

N/Alternate names: cysteine proteinase inhibitor; onchocerciasis antigen
 C/Species: Onchocerca volvulus
 C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
 C/Accession: A43428; B43927
 R/Lustigman, S.; Brozman, B.; Huima, T.; Prince, A.M.; Mckerrow, J.H.
 J. Biol. Chem. 267, 17339-17346, 1992
 A/Title: Molecular cloning and characterization of onchocystatin, a cysteine proteinase
 A/Reference number: A43428; MUID:92381053; PMID:1512269
 A/Accession: A43428
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-162 <IUS>
 A/Cross-references: GB:U37105; NID:g159905; PID:g159906
 A/Note: sequence extracted from NCBI backbone (NCBIN:111962, NCBIPI:111963)
 R/Chandrahekar, R.; Masood, K.; Alvarez, R.M.; Ogunrinade, A.F.; Lujan, R.; Richards Jr
 J. Clin. Invest. 88, 1460-1466, 1991
 A/Title: Molecular cloning and characterization of recombinant parasite antigens for imm
 A/Reference number: A43927; MUID:92042723; PMID:1840605
 A/Accession: B43927
 A/Molecule type: mRNA
 A/Residues: 'P', '37-57', 'A', '59-71', 'R', '73-83', 'N', '85-126', 'W', '128-162' <CHA>
 A/Cross-references: GB:M60279; NID:g159888
 A/Experimental source: clone OC 9.3
 A/Note: sequence extracted from NCBI backbone (NCBIN:65111, NCBIPI:65113)
 C/Superfamily: cystatin; cystatin homology
 C/Keywords: cysteine proteinase inhibitor

Query Match 39.2%; Score 56; DB 2; Length 162;
 Best Local Similarity 45.5%; Pred. No. 0.51;
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 4 NKESDDKXFRFRFLVKYQROV 25
 DB 77 NEQSNDEYHMPKILKVSQV 98

RESULT 12

UDHUP1

cystatin S precursor - human

N/Alternate names: cystatin SA-II; salivary acidic protein-1

C/Species: Homo sapiens (man)

C/Date: 25-Feb-1985 #sequence_revision 08-Feb-1996 #text_change 16-Jul-1999

C/Accession: S17667; S16500; A01272; A29603; S19280; A56608

R/Bober, L.A.; Aguirre, A.; Levine, M.J.

Biochem. J. 278, 627-635, 1991

A/Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ an

A/Reference number: S17667; MUID:91378918; PMID:1898352

A/Accession: S17667

A/Molecule type: mRNA

A/Residues: 1-141 <BOB>

A/Cross-references: EMBL:X54667; NID:930365; PIDN:CA38478.1; PID:930366

R/Lamkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.

Arch. Biochem. Biophys. 288, 664-670, 1991

A/Title: Salivary cystatin SA-II, a potential precursor of the acquired enamel pellicle

A/Reference number: S16500; MUID:91378515; PMID:1898055

A/Accession: S16500

A/Status: preliminary

A/Molecule type: protein

A/Residues: 21-134, 'D', '136-141' <IHU>

R/Isemura, S.; Saitoh, E.; Sanada, K.

J. Biochem. 96, 489-498, 1984

A/Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal

A/Reference number: A91985; MUID:85054716; PMID:6501254

A/Accession: A01272

A/Molecule type: protein

A/Residues: 29-134, 'D', '136-141' <ISE>

R/Isemura, S.; Saitoh, E.; Ito, S.; Isemura, M.; Sanada, K.

J. Biochem. 96, 1311-1314, 1984

A/Title: Cystatin S: a cysteine proteinase inhibitor of human saliva.

A/Reference number: A91981; MUID:85104877; PMID:6394600

A/Contents: annotation; inhibitor specificity

R/Hawke, D.H.; Yuan, P.M.; Wilson, K.J.; Hunkapiller, M.W.

Biochem. Biophys. Res. Commun. 145, 1248-1253, 1987

A/Title: Identification of a long form of cystatin from human saliva by rapid microbo
 A/Reference number: A29603; MUID:87270697; PMID:3496880
 A/Accession: A29603
 A/Molecule type: protein
 A/Residues: 21-51 <HAW>
 R/Ramasubbu, N.; Reddy, M.S.; Bergey, E.J.; Haraazthy, G.G.; Soni, S.D.; Levine, M.J.
 Biochem. J. 280, 341-352, 1991
 A/Title: Large-scale purification and characterization of the major phosphoproteins and
 A/Reference number: S19279; MUID:92082469; PMID:1747107
 A/Accession: S19280
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 21-55 <RAM>
 R/Johnson, M.; Richardson, C.F.; Bergey, E.J.; Levine, M.J.; Nancollas, G.H.
 Arch. Oral Biol. 36, 631-636, 1991
 A/Title: The effects of human salivary cystatin and statherin on hydroxyapatite crystal
 A/Reference number: A56608; MUID:92074899; PMID:1741653
 A/Accession: A56608
 A/Molecule type: protein
 A/Residues: 21-36 <JOH>
 A/Note: sequence extracted from NCBI backbone (NCBIPI:67866)
 C/Comment: This protein strongly inhibits papain and ficin, partially inhibits stem brom
 competitively.

A/Genes: GDB:CST4
 A/Cross-references: GDB:136381

A/Map position: 20p11.2-20p11.2

C/Superfamily: cystatin; cystatin homology

C/Keywords: cysteine proteinase inhibitor; phosphoprotein; saliva

F/1-20/Domain: signal sequence #status predicted <SIG>

F/21-141/Product: cystatin S #status predicted <MAT>

F/30-141/Domain: cystatin homology <CYS>

F/76-80/Region: inhibitory #status predicted

F/94-104,118-138/Disulfide bonds: #status predicted

Query Match 37.8%; Score 54; DB 1; Length 141;
 Best Local Similarity 39.1%; Pred. No. 0.88;
 Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 2 QYNKESDDKXFRFRFLVKYQRO 24
 DB 54 EYNKATEDEYRRLPVLRLAREQ 76

RESULT 13

UDHUP2

cystatin SN precursor [validated] - human

N/Alternate names: cystatin SA-I

C/Species: Homo sapiens (man)

C/Date: 28-May-1986 #sequence_revision 08-Feb-1996 #text_change 08-Dec-2000

C/Accession: A28110; S02489; A29632; A01273; S19279

R/Al-Hachimi, I.; Dickinson, D.P.; Levine, M.J.

J. Biol. Chem. 263, 9381-9387, 1988

A/Title: Purification, molecular cloning, and sequencing of salivary cystatin SA-I.

A/Reference number: A28110; MUID:88243825; PMID:2837486

A/Accession: A28110

A/Molecule type: mRNA

A/Residues: 1-141 <ALH>

A/Cross-references: GB:U03870; NID:9337751; PIDN:AAA60299.1; PID:9337752

R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A/Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A/Reference number: S02489; MUID:89076505; PMID:3202964

A/Accession: S02489

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 21-141 <SA2>

R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A/Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three memt

A/Reference number: A91589; MUID:88185836; PMID:3446578

A/Accession: A29632

A:Molecule type: protein

A:Residues: 1-120 <ESN>

A>Note: 43.Asn was also found

A>Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala

C:Superfamily: cystatin; cystatin homology <CYS>

F:9-120/Domain: cystatin homology <CYS>

Query Match 43.4%; Score 62; DB 2; Length 120;

Best Local Similarity 41.7%; Pred. No. 0.047;

Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

2 QYNKESDDKHYFRIFRYLKVGROV 25

33 EYKGSNDAYHSRAIQVVRARKQL 56

RESULT 7

S07085

Cystatin C precursor - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 01-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999

C:Accession: S07085; S01337; S21109

R:Cole, T.; Dickson, P.W.; Esnard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Sch

Eur. J. Biochem. 186, 35-42, 1989

A>Title: The cDNA structure and expression analysis of the genes for the cysteine prote

A:Reference number: S07085; MUID:90092122; PMID:12689174

A:Accession: S07085

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-127 <COL>

A:Cross-references: EMBL:X16957; NID:956041; PIDN:CAA34831.1; PID:G736290

R:Esnard, A.; Esnard, F.; Faucher, D.; Gauthier, F.

FEBS Lett. 236, 475-478, 1988

A>Title: Two rat homologues of human cystatin C.

A:Reference number: S01337; MUID:88313020; PMID:3044831

A:Accession: S01337

A:Molecule type: protein

A:Residues: 8-49 <ESN>

R:Esnard, A.; Esnard, F.; Guillon, F.; Gauthier, F.

FEBS Lett. 300, 131-135, 1992

A>Title: Production of the cysteine proteinase inhibitor cystatin C by rat Sertoli cells

A:Reference number: S21109; MUID:9225121; PMID:1563513

A:Accession: S21109

A:Molecule type: protein

A:Residues: 8, 'XX', 11-20 <ES2>

C:Superfamily: cystatin; cystatin homology

C:Keywords: cysteine proteinase inhibitor

F:16-127/Domain: cystatin homology <CYS>

F:80-90,104-124/Disulfide bonds: #status predicted

Query Match 43.4%; Score 62; DB 2; Length 127;

Best Local Similarity 41.7%; Pred. No. 0.05;

Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

2 QYNKESDDKHYFRIFRYLKVGROV 25

40 EYKGSNDAYHSRAIQVVRARKQL 63

RESULT 8

A36163

Cystatin C precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 16-Jul-1999

C:Accession: A36163

R:Solem, M.; Rawson, C.; Lindburg, K.; Barnes, D.

Biochem. Biophys. Res. Commun. 172, 945-951, 1990

A>Title: Transforming growth factor beta regulates cystatin C in serum-free mouse embryo

A:Reference number: A36163; MUID:91054522; PMID:2241983

A:Accession: A36163

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <SOL>

A:Cross-references: EMBL:M59470; NID:G192911; PIDN:AAA63298.1; PID:G192912

C:Superfamily: cystatin; cystatin homology

F:28-140/Domain: cystatin homology <CYS>

F:93-103,117-137/Disulfide bonds: #status predicted

Query Match 43.4%; Score 62; DB 2; Length 140;

Best Local Similarity 41.7%; Pred. No. 0.055;

Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

2 QYNKESDDKHYFRIFRYLKVGROV 25

53 EYKGSNDAYHSRAIQVVRARKQL 76

RESULT 9

E64005

hypothetical protein HI0284 - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999

C:Accession: E64005

R:Reichmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A

; Goeyne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glöck, A.; Kelley, J.M.; Weidman, J

; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Georgagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: E64005

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-139 <TRGR>

A:Cross-references: GB:U32714; GB:U42023; NID:G1573241; PID:G1573253; TIGR:HI0284

C:Genetic

A:Start codon: GTG

C:Superfamily: Haemophilus influenzae hypothetical protein HI0284

Query Match 41.3%; Score 59; DB 2; Length 139;

Best Local Similarity 45.0%; Pred. No. 0.15;

Matches 9; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

7 SDDKHYFRIFRYLKVGROV 26

91 NDKKFFYKIPISIKIQKRLT 110

RESULT 10

S62326

Cystatin C, parotid glands - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999

C:Accession: S62326

R:Ciemerman, N.; Drobnic Kosorok, M.; Korant, B.D.; Turk, B.; Turk, V.

Biol. Chem. Hoppe-Seyler 377, 19-23, 1996

A>Title: Characterization of cystatin C from bovine parotid glands: cysteine proteinase

A:Reference number: S62326; MUID:97083236; PMID:8929810

A:Accession: S62326

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-47 <CIM>

C:Superfamily: cystatin; cystatin homology

Query Match 40.6%; Score 58; DB 2; Length 47;

Best Local Similarity 39.1%; Pred. No. 0.068;

Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

2 QYNKESDDKHYFRIFRYLKVGROQ 24

25 EFKKSNDAQSRVVRVVRARKQ 47

RESULT 11

A43428

onchocystatin - nematode (Onchocerca volvulus)

A:Reference number: S55305; MUID:94379969; PMID:8092991
A:Accession: S55305
A>Status: preliminary
A:Molecule type: protein
A:Residues: 27-49;106-146 <BBR>
C:Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fluid from patients with certain autoimmune diseases.
C:Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorrhage in Danish kindreds:
C:Gene: GDB:CST3
A:Cross-references: GDB:119817; OMIM:105150
A:Map position: 20p11.2-20p11.2
A:Ions: 81/3; 119/3
C:Superfamily: cystatin; cystatin homology
C:Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline
F:1-66/Domain: signal sequence #status predicted <Sig>
F:127-146/Product: cystatin C #status experimental <Mat>
F:335-146/Domain: cystatin homology <Cys>
E:81-85/Region: inhibitory #status predicted
F:29/Modified site: hydroxyproline (Pro) (partial) #status experimental
F:99-109;123-143/Disulfide bonds: #status experimental

Query Match 45.5% Score 65; DB 1; Length 146;
Best Local Similarity 41.7%; Pred. No. 0.021;
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRIFRLVKVOROV 25
 :|||:::||:||||:~::~||
Db 59 EYNKASNDMYHSRALQVVVRARQOI 82

RESULT 5

UDBO
cystatin - bovine
N:Alternate names: thiol proteinase inhibitor
C:Species: Bos primigenius taurus (catle)
C>Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #ext_change 06-Dec-1996
C:Accession: A01271
S:Hirado, M.; Tsunashima, S.; Sakiyama, F.; Ninobe, M.; Fujii, S.
FEBS Lett. 186, 41-45, 1985
A>Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase inhibitor
A:Reference number: A01271; MUID:85321205; PMID:3891407
A:Accession: A01271
A:Molecule type: Protein
A:Residues: 1-112 <HIR>
C:Superfamily: cystatin; cystatin homology
C:Keywords: colostrum; cysteine proteinase inhibitor
F:2-112/Domain: cystatin homology <CYS>
F:48-52/Region: inhibitory #status predicted
F:66-76;90-110/Dissulfide bonds: #status predicted

Query Match 43.4% Score 62; DB 1; Length 112;
Best Local Similarity 41.7%; Pred. No. 0.043;
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRIFRLVKVOROV 25
 :|||:::||:||||:~::~||
Db 26 EFNRKSNDAVSRRVVVRARQOV 49

RESULT 6

cystatin C - rat
C:Species: Rattus sp. (rat)
C>Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #ext_change 16-Jul-1999
C:Accession: S10587
S:Esnard, F.; Esnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gautahier, B.; Hoppé-Seyler 371(Suppl.), 161-166, 1990
BioJl. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990
A>Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation
A:Reference number: S10587; MUID:9080276; PMID:2400577
A:Accession: S10587
A>Status: preliminary

A>Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cytatatin
A:Reference number: S01461; MUID:84110059; PMID:6662498

A:Accession: S01461

A:Molecule type: protein

A:Residues: 24-139 <TUR>

R:Amanstetel, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.
Biochem J. 211, 129-138, 1983

A>Title: Cystatatin, a protein inhibitor of cysteine proteenases. Improved purification frt
A:Reference number: A57514; MUID:83256421; PMID:6403085

A:Contents: annotation; characterization of protein

R:Grubb, A., Lofberg, H., Barrett, A.J.
FEBS Lett. 170, 370-374, 1984

A>Title: The disulfide bridges of human cystatatin C (gamma-trace) and chicken cystatatin.

A:Reference number: S01462

A:Contents: annotation; disulfide bonds

R:Auerwald, E.A.; Naegler, D.K.; Schulze, A.J.; Engh, R.A.; Genenger, G.; Machleidt, W.
Eur J Biochem. 224, 407-415, 1994

A>Title: Production, inhibitory activity, folding and conformational analysis of an N-te
A:Reference number: 548159; MUID:95010016; PMID:7925354

A:Accession: 548159

A>Status: preliminary

A:Molecule type: protein

A:Residues: 24-139 <AUE>

R:Labeer, B.; Kriesling, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.
FEBS Lett. 248, 162-168, 1989

A>Title: The cysteine proteinase inhibitor chicken cystatatin is a phosphoprotein.

A:Reference number: S04008; MUID:89252033; PMID:2721673

A:Accession: S04008

A:Molecule type: protein

A:Residues: 97-114 <LAB>

R:Cotella, R.; Bird, T.W.C.
Gene 130, 175-181, 1993

A>Title: Isolation and characterization of the chicken cystatatin-encoding gene: Mapping t
A:Reference number: UN0789; MUID:93366172; PMID:8359684

A:Accession: UN0789

A:Molecule type: DNA

A:Residues: 1-139 <CO2>

A:Cross-references: GS:M95725

A>Note: authors failed to translate the codon for residue 115-Tyr

C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases

C:Genetics:

A:Gene: Cam

A:Introns: 76/3; 114/3

C:Superfamily: cystatatin; cystatatin homology

C:Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein

F:1-23/Domian: signal sequence #status predicted <SIG>

F:24-139/Product: cystatatin, long form #status experimental <CLF>

E:30-139/Domain: cystatatin homology <CYS>

F:32-139/Product: cystatatin, short form #status experimental <CYSP>

F:76-80/Region: inhibitory #status predicted

F:94-104,118-138/disulfide bonds: #status experimental

F:103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

A: Molecule type: DNA
A: Residues: 1-141 <SA1>
A: Cross-references: GB:M19673; GB:M19170; NID:9186403; PIDN:AAA6116.1; PID:9386826
A: Note: The authors translated the codon GAC for residue 129 as Asn
R: Satoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988
A: Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily
A: Reference number: S02489; MUID:89076505; PMID:3202964
A: Accession: S02490
A: Status: not compared with conceptual translation
A: Molecule type: DNA
A: Residues: 21-141 <SA2>
R: Isemura, S.; Satoh, E.; Sanada, K.
J. Biochem. 102, 693-704, 1987
A: Title: Characterization and amino acid sequence of a new acidic cysteine proteinase from
A: Reference number: A41422; MUID:88139220; PMID:3436950
A: Accession: A41422
A: Molecule type: protein
A: Residues: 25-141 <IS>
R: Isemura, S.; Satoh, E.; Sanada, K.; Isemura, M.; Ito, S.
in Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruyter
A: Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.
A: Reference number: A27015
A: Accession: B27015
A: Molecule type: protein
A: Residues: 25-134, 'D', 136-141 <IS2>
C: Genetics:
A: Gene: GDB: CST2
A: Cross-references: GDB:119816; OMIM:123856
A: Map position: 20p11.2-20p11.2
C: Superfamily: cystatin; cystatin homology
F: 30-141/Domain: cystatin homology <CS>

QY 2 OYNKESDDKYHRIPIVLTAKVOROV 25
:|::||::|::|::|::|:
Db 54 EYNKATEDEYYRRLRVLPAREOI 77

RESULT 4

UDHU

Cystatin C precursor [validated] - human

M.Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma I

C.Species: Homo sapiens (man)

C.Date: 06-Jul-1982 #sequence_revision 13-Mar-1991 #text_change 08-Dec-2000

C.Accession: S10216; S000004; J100095; A33440; S02751; A01270; A25434; S12288; A32732; A601

R.Abrahamson, M.; Olafsson, I.; Paladottir, A.; Ulvsbeck, M.; Lundwall, A.; Jensen, O.,

Biochem. J. 268, 287-294, 1990

A.Title: Structure and expression of the human cystatin C gene.

A.Reference number: S10216; MWID:90303202; PMID:2363674

A.Accession: S10216

A.Molecule type: DNA

A.Residues: 1-146 <AB1>

A.Cross-references: EMBL:X52255; NID:g30257; PIDN:CMA36497.1; PID:g296643

R.Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FBS Lett. 216, 229-233, 1987

A>Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A.Reference number: S00004; MWID:87219149; PMID:3395457

A.Accession: S00004

A.Molecule type: mRNA

A.Residues: 1-146 <AB2>

A.Cross-references: EMBL:X05607; NID:g30371; PIDN:CMA29096.1; PID:g755738

R.Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.

J. Exp. Med. 169, 1771-1778, 1989

A>Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A.Reference number: J10095; MWID:89235594; PMID:2541223

A.Accession: J10095

A.Molecule type: DNA

A.Residues: 1-146 <LEV>

A.Cross-references: GB:X61681; NID:g30367; PIDN:CMA43856.2; PID:g4450944

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 5.19643 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-9
Perfect score: 143
Sequence: 1 DYNKESDDKYHFRFVLKVKORQVTD 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	76	53.1	139	2	A45361	cystatin-related e
2	66	46.2	139	1	UDCH	cystatin precursor
3	66	46.2	141	2	B29632	cystatin SA precu
4	65	45.5	146	1	UDHU	cystatin C precu
5	62	43.4	112	1	UDBO	cystatin - bovine
6	62	43.4	120	2	S10587	cystatin C - rat
7	62	43.4	127	2	S07085	cystatin C precu
8	62	43.4	140	2	A36163	cystatin C precu
9	59	41.3	139	2	B64005	hypothetical prote
10	58	40.6	47	2	B62326	cystatin C, parot
11	56	39.2	162	2	A43428	onchocystatin - ne
12	54	37.8	141	1	UDHUP1	cystatin S precu
13	54	37.8	141	1	UDHUP2	cystatin SN precu
14	53	37.1	141	2	Q01470	cystatin S precu
15	51	35.7	111	2	A28793	cystatin - puff ad
16	50.5	35.3	394	2	T33641	hypothetical prote
17	50	35.0	319	2	A86650	riamosyltransfera
18	50	35.0	393	2	D86168	hypothetical prote
19	49	34.3	442	2	T06640	hypothetical prote
20	49	34.3	498	2	T31871	hypothetical prote
21	49	34.3	525	1	KGHUGH	histidine-rich gly
22	49	34.3	1146	2	S46837	hypothetical prote
23	48	33.6	917	1	RDMUNH	nitrate reductase
24	47.5	33.2	505	1	PSXR44	outer capsid prote
25	47	32.9	224	2	S54519	HuJ1 protein - yea
26	47	32.9	370	2	T47762	hypothetical prote
27	47	32.9	602	2	S69198	prostaglandin G/H
28	47	32.9	602	2	S39782	cyclooxygenase 1 -
29	47	32.9	602	2	A35564	prostaglandin-endo

30	47	32.9	681	2	T01469	hypothetical prote
31	47	32.9	1973	2	G89608	protein B0272.5 [1
32	47	32.9	1973	2	T18686	hypothetical prote
33	47	32.9	2305	2	T15571	hypothetical prote
34	47	32.9	2305	2	B89608	protein C23P12.1 [
35	46	32.2	533	2	S57904	virR49 protein - S
36	46	32.2	533	2	UC5040	positive regulator
37	46	32.2	656	2	T52064	dnaJ-like protein
38	46	32.2	734	2	S37998	probable serine/th
39	45.5	31.8	160	2	B82710	conserved hypothet
40	45.5	31.8	262	2	S58288	heterocyst differe
41	45.5	31.8	273	2	T17629	hypothetical prote
42	45.5	31.8	430	2	B69620	phosphopyruvate hy
43	45.5	31.8	916	2	AG1902	acetyl-CoA synchet
44	45	31.5	73	2	Q01921	hypothetical 8.5k
45	45	31.5	111	1	JC2040	cystatin - chum sa

ALIGNMENTS

RESULT 1
A45361
cystatin-related epididymal specific protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C/Accession: A45361
R/Cornwall, G.A.; Orgebin-Crist, M.C.; Hamu, S.R.
Mol. Endocrinol. 6, 1653-1664, 1992
A/Title: The CRIS gene: a unique testis-regulated gene related to the cystatin family is
A/Reference number: A45361; MUID:93078799; PMID:1280328
A/Accession: A45361
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-139 <COR>
A/Cross-references: GB:S49926; NID:g260492; PIRN:AAC35390.1; PID:g260493
A/Note: sequence extracted from NCBI backbone (NCBIP:118813)
C/Superfamily: cystatin; cystatin homology
F/28-139/Domain: cystatin homology <CYS>

Query Match 53.1%; Score 76; DB 2; Length 139;
Best Local Similarity 50.0%; Pred. No. 0.00044;
Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
QY 2 DYNKESDDKYHFRFVLKVKORQVTD 27
DB 52 EYNKESDDKYHFRFVLKVKORQVTD 77
UDCH
cystatin precursor - chicken
N/Alternate names: cystatin 1; cysteine proteinase inhibitor; egg-white cystatin
C/Species: Gallus gallus (chicken)
C/Date: 03-Aug-1984 #sequence_revision 12-Apr-1996 #text_change 29-Oct-1999
C/Accession: A4456; A01274; S01461; S48159; S04008; JN0789
R/Colletta, R.; Sakaguchi, Y.; Nagase, H.; Bird, J.W.C.
J. Biol. Chem. 264, 17154-17169, 1989
A/Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue c
A/Reference number: A4456; MUID:9008873; PMID:2793849
A/Accession: A4456
A/Molecule type: mRNA
A/Residues: 1-139 <COL>
A/Cross-references: GB:Y05077; NID:g211714; PIRN:AAA46744.1; PID:g211715
R/Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.K.; Barrett, A.U.
Biochem. J. 217, 813-817, 1984
A/Title: Cystatin. Amino acid sequence and possible secondary structure.
A/Reference number: A01274; MUID:84178305; PMID:6712597
A/Accession: A01274
A/Molecule type: protein
A/Residues: 24-139 <SCH>
R/Turk, V.; Bzian, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983

THIS PAGE BLANK (USPTO)

Query Match 46.2%; Score 66; DB 9; Length 116;
 Best Local Similarity 45.8%; Pred. No. 0.038;
 Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKXQROV 25
 :||:|||||:|:|:|:
 Db 31 EYNRASNDKYSRVRVISAQRDL 54

RESULT 14

US-09-775-932-8
 ; Sequence 8, Application US/09775932
 ; Patent No. US20020137671A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of British Columbia
 ; TITLE OF INVENTION: Production and use of Modified Cystatins
 ; FILE REFERENCE: 58069
 ; CURRENT APPLICATION NUMBER: US/09/775,932
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: CA99/00717
 ; PRIOR FILING DATE: 1999-08-05
 ; PRIOR APPLICATION NUMBER: 60/095,503
 ; PRIOR FILING DATE: 1998-08-05
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 121
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-775-932-8

Query Match 46.2%; Score 66; DB 9; Length 121;
 Best Local Similarity 41.7%; Pred. No. 0.04;
 Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKXQROV 25
 :||:|||||:|:|:|:
 Db 34 EYNKATEDEYRRLRLVLRAREQI 57

RESULT 15

US-08-849-303-15
 ; Sequence 15, Application US/08849303
 ; Publication No. US20030221209A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Atkinson, Howard J.
 ; APPLICANT: McPherson, Michael J.
 ; APPLICANT: Utwin, Peter E.
 ; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
 ; NUMBER OF SEQUENCES: 79
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue, 4th Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/849,303
 ; FILING DATE: 21-MAY-1997
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Bag., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 1321-1-003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684

TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 139 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; US-08-849-303-15

Query Match 46.2%; Score 66; DB 8; Length 139;
 Best Local Similarity 45.8%; Pred. No. 0.046;
 Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKXQROV 25
 :||:|||||:|:|:|:
 Db 54 EYNRASNDKYSRVRVISAQRDL 77

Search completed: March 18, 2004, 14:30:53
 Job time : 16.3393 secs

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48936
LENGTH: 50
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096677.18
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
OTHER INFORMATION: EST HUMAN HIT: A120857.1, EVALU 5.00e-23
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALU 1.00e-01
US-09-864-761-48936

Query Match 89.5%; Score 128; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKYHFRIFRLVKVQROV 24
Db 27 DQNKESDDKYHFRIFRLVKVQROV 50

RESULT 10
US-09-941-314-10
Sequence 10, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: Zymogenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-10

Query Match 87.4%; Score 125; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NKESDDKYHFRIFRLVKVQROVTD 27
Db 1 NKESDDKYHFRIFRLVKVQROVTD 24

RESULT 11
US-09-941-314-14
Sequence 14, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: Zymogenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314

CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 52
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-14

Query Match 83.2%; Score 119; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KESDDKYHFRIFRLVKVQROVTD 27
Db 1 KESDDKYHFRIFRLVKVQROVTD 23

RESULT 12
US-10-264-049-2608
Sequence 2608, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P4133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2608
LENGTH: 181
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2608

Query Match 51.7%; Score 74; DB 15; Length 181;
Best Local Similarity 54.2%; Pred. No. 0.004;
Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRIFRLVKVQROV 25
Db 94 EYNKASNDAYHSRAVRVVRKQV 117

RESULT 13
US-09-775-932-16
Sequence 16, Application US/09775932
Patent No. US20020137671A1
GENERAL INFORMATION:
APPLICANT: University of British Columbia
TITLE OF INVENTION: Production and use of Modified Cystatins
FILE REFERENCE: 58069
CURRENT APPLICATION NUMBER: US/09/775,932
CURRENT FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: CA99/00717
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: 60/095,503
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 116
TYPE: PRT
ORGANISM: Gallus sp.
US-09-775-932-16

GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 80
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-15

Query Match 95.8%; Score 137; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIYLVKQROVTD 27
Db 1 QYNKESDDKXHFRIYLVKQROVTD 26

RESULT 8
US-09-864-761-34822
Sequence 34822, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34822
LENGTH: 50
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL109954.10
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match 89.5%; Score 128; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIYLVKQRO 24
Db 27 DQYNKESDDKXHFRIYLVKQRO 50

RESULT 9
US-09-864-761-48936
Sequence 48936, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408

FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29,230
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 35
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-8

Query Match 100.0%; Score 143; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 4,2e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFFVLKVQRQVTD 27
Db 9 DQYNKESDDKXHFRIFFVLKVQRQVTD 35

RESULT 3
US-09-941-314-13
Sequence 13, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29,230
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 49
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-13

Query Match 100.0%; Score 143; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 6,1e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFFVLKVQRQVTD 27
Db 1 DQYNKESDDKXHFRIFFVLKVQRQVTD 27

RESULT 4
US-09-941-314-3
Sequence 3, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29,230
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-3

Query Match 100.0%; Score 143; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 1,5e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFFVLKVQRQVTD 27
Db 30 DQYNKESDDKXHFRIFFVLKVQRQVTD 56

RESULT 5
US-09-941-314-4
Sequence 4, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29,230
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-4

Query Match 100.0%; Score 143; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 1,6e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFFVLKVQRQVTD 27
Db 32 DQYNKESDDKXHFRIFFVLKVQRQVTD 58

RESULT 6
US-09-941-314-2
Sequence 2, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29,230
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 137
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-2

Query Match 100.0%; Score 143; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 1,8e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKXHFRIFFVLKVQRQVTD 27
Db 52 DQYNKESDDKXHFRIFFVLKVQRQVTD 78

RESULT 7
US-09-941-314-15
Sequence 15, Application US/09941314
Patent No. US20020142396A1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 16.3393 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-9
Perfect score: 143
Sequence: 1 DQYNKESDDKYHPRFRVLKVRQVTD 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	27	US-09-941-314-9	Sequence 9, Appli
2	143	100.0	35	US-09-941-314-8	Sequence 8, Appli
3	143	100.0	49	US-09-941-314-13	Sequence 13, Appli
4	143	100.0	115	US-09-941-314-3	Sequence 3, Appli
5	143	100.0	117	US-09-941-314-4	Sequence 4, Appli
6	143	100.0	137	US-09-941-314-2	Sequence 2, Appli
7	137	95.8	80	US-09-941-314-15	Sequence 15, Appli
8	128	89.5	50	US-09-864-761-34822	Sequence 34822, A
9	125	87.4	46	US-09-864-761-48936	Sequence 48936, A
10	119	83.2	52	US-09-941-314-10	Sequence 10, Appli
11	119	83.2	52	US-09-941-314-14	Sequence 14, Appli
12	74	51.7	181	US-10-264-049-2608	Sequence 2608, Ap
13	66	46.2	116	US-09-775-932-16	Sequence 16, Appli
14	66	46.2	121	US-09-775-932-8	Sequence 8, Appli
15	66	46.2	139	US-08-849-303-15	Sequence 15, Appli

16	66	46.2	139	9	US-09-969-834-4	Sequence 4, Appli
17	66	46.2	141	8	US-08-849-303-24	Sequence 24, Appli
18	66	46.2	141	9	US-09-940-497-6	Sequence 6, Appli
19	65	45.5	120	9	US-09-775-932-2	Sequence 2, Appli
20	65	45.5	146	8	US-08-849-303-17	Sequence 17, Appli
21	65	45.5	146	9	US-09-940-497-3	Sequence 3, Appli
22	65	45.5	146	9	US-09-969-834-3	Sequence 3, Appli
23	65	45.5	146	14	US-10-329-428-3	Sequence 3, Appli
24	65	45.5	146	14	US-10-376-564-47	Sequence 47, Appli
25	63	44.1	145	9	US-09-740-638-2	Sequence 2, Appli
26	63	44.1	145	13	US-10-006-467-2	Sequence 2, Appli
27	63	44.1	145	14	US-10-235-148-2	Sequence 2, Appli
28	62	43.4	112	8	US-08-849-303-16	Sequence 16, Appli
29	62	43.4	118	9	US-09-775-932-24	Sequence 24, Appli
30	62	43.4	127	8	US-08-849-303-19	Sequence 19, Appli
31	62	43.4	140	8	US-08-849-303-18	Sequence 18, Appli
32	62	43.4	140	14	US-10-376-564-46	Sequence 46, Appli
33	62	43.4	140	14	US-10-376-564-48	Sequence 48, Appli
34	61	42.7	141	9	US-09-775-932-18	Sequence 18, Appli
35	57	39.9	145	14	US-10-168-425-14	Sequence 14, Appli
36	56	39.2	162	9	US-08-849-303-25	Sequence 25, Appli
37	55	38.5	18	9	US-09-941-314-7	Sequence 7, Appli
38	55	38.5	36	9	US-09-941-314-6	Sequence 6, Appli
39	54	37.8	121	9	US-09-775-932-4	Sequence 4, Appli
40	54	37.8	121	9	US-09-775-932-6	Sequence 6, Appli
41	54	37.8	141	8	US-08-849-303-21	Sequence 21, Appli
42	54	37.8	141	8	US-08-849-303-22	Sequence 22, Appli
43	54	37.8	141	9	US-09-940-497-5	Sequence 5, Appli
44	54	37.8	141	9	US-09-940-497-7	Sequence 7, Appli
45	54	37.8	141	9	US-09-974-298-141	Sequence 141, App

ALIGNMENTS

RESULT 1
US-09-941-314-9
; Sequence 9, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-9

Query Match 100.0%; Score 143; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.2e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYNKESDDKYHPRFRVLKVRQVTD 27
Db 1 DQYNKESDDKYHPRFRVLKVRQVTD 27

RESULT 2
US-09-941-314-8
; Sequence 8, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

THIS PAGE BLANK (USPTO)

```
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Human Cystatin E
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,030
FILING DATE: 05-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PP202PID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301 309 8504
TELEFAX: 301 309 8512
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Cystatin SN
US-09-241-376-6

Query Match 46.2%; Score 66; DB 4; Length 141;
Best Local Similarity 41.7%; Pred. No. 0.0045;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKYQROV 25
Db 54 EYNKATEDERYRRLRLVLRARBOI 77

RESULT 14
US-09-940-497-6
Sequence 6, Application US/09940497
Patent No. 6617132
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: Human Cystatin E
FILE REFERENCE: PP202PID2
CURRENT APPLICATION NUMBER: US/09/940,497
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: US 09/241,376
PRIOR FILING DATE: 1999-02-02
PRIOR APPLICATION NUMBER: US 08/744,138
PRIOR FILING DATE: 1996-11-05
PRIOR APPLICATION NUMBER: US 08/461,030
PRIOR FILING DATE: 1995-06-05
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.1
```

```
SEQ ID NO 6
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-940-497-6
```

```
Query Match 46.2%; Score 66; DB 4; Length 141;
Best Local Similarity 41.7%; Pred. No. 0.0045;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 2 QYNKESDXYHFRIFRYLKYQROV 25
Db 54 EYNKATEDERYRRLRLVLRARBOI 77
```

RESULT 15

```
US-08-849-303-24
Sequence 24, Application US/08849303
Patent No. 6680424
GENERAL INFORMATION:
APPLICANT: Atkinson, Howard J.
APPLICANT: McPherson, Michael J.
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TEXT: 133521
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-24
```

```
Query Match 46.2%; Score 66; DB 4; Length 141;
Best Local Similarity 41.7%; Pred. No. 0.0045;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 2 QYNKESDXYHFRIFRYLKYQROV 25
Db 54 EYNKATEDERYRRLRLVLRARBOI 77
```

```
Search completed: March 18, 2004, 14:25:36
Job time : 6.91071 secs
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/791,522
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0193 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 118195
US-09-314-777-4

Query Match 46.2%; Score 66; DB 3; Length 139;
Best Local Similarity 45.8%; Pred. No. 0.0044;
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 2 QYNKESDXYHFRIPRLKXQROV 25
Db 54 EYNRASNDKXSRVVRVISAQR 77

RESULT 11
US-08-849-303-15
Sequence 15, Application US/08849303
Patent No. 6680424
GENERAL INFORMATION:
APPLICANT: Atkinson, Howard J.
APPLICANT: McPherson, Michael J.
APPLICANT: Urvn, Peter E.
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-15

Query Match 46.2%; Score 66; DB 4; Length 139;
Best Local Similarity 45.8%; Pred. No. 0.0044;
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 2 QYNKESDXYHFRIPRLKXQROV 25
Db 54 EYNRASNDKXSRVVRVISAQR 77

RESULT 12
US-08-744-138-6
Sequence 6, Application US/08744138
Patent No. 6011012
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner L.
APPLICANT: Ni, Jian
APPLICANT: Rosen, Craig A.
APPLICANT: Yu, Guo-liang
TITLE OF INVENTION: Human Cystatin B
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/744,138
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF202P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301 309 8504
TELEFAX: 301 309 8512
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Cystatin SN
US-08-744-138-6

Query Match 46.2%; Score 66; DB 3; Length 141;
Best Local Similarity 41.7%; Pred. No. 0.0045;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QYNKESDXYHFRIPRLKXQROV 25
Db 54 EYNKATEDYRKLRLVLRARBOI 77

RESULT 13
US-09-241-376-6
Sequence 6, Application US/09241376
Patent No. 6306477
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner L.
APPLICANT: Ni, Jian
APPLICANT: Rosen, Craig A.

```
RESULT 7
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cycatans
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match          46.2%; Score 66; DB 4; Length 116;
Best Local Similarity 45.8%; Pred. No. 0.0036;
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      2 QYNKESDXYHFRIFRLVKYQROV 25
DB      31 EYNRASNDKYSRVVRVISAQR 54

RESULT 8
US-09-775-932-8
; Sequence 8, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cycatans
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-775-932-8

Query Match          46.2%; Score 66; DB 4; Length 121;
Best Local Similarity 41.7%; Pred. No. 0.0037;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY      2 QYNKESDXYHFRIFRLVKYQROV 25
DB      34 EYNKATEDEYRRLRLVLRARAEQ 57

RESULT 9
US-08-791-522-4
; Sequence 4, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bardman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
```

```
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,522
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0193 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 118195
US-08-791-522-4

Query Match          46.2%; Score 66; DB 2; Length 139;
Best Local Similarity 45.8%; Pred. No. 0.0044;
Matches 11; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      2 QYNKESDXYHFRIFRLVKYQROV 25
DB      54 EYNRASNDKYSRVVRVISAQR 77

RESULT 10
US-09-314-777-4
; Sequence 4, Application US/09314777
; Patent No. 610686
; GENERAL INFORMATION:
; APPLICANT: Bardman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,777
; FILING DATE:
```

```

; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-617-302-2

```

Query Match	56.6%	Score 81	DB 3	Length 141
Best Local Similarity	53.8%	Pred. No.	1.9e-05	
Matches 14	Conservative	6	Mismatches 6	Indels 0
			Gaps	0

```

QY      2 QYNKESDDKTHFRIFRLVLKVRQVTD 27
          :|||:|:|:|:|:|:|:|:|:|:|
DB      54 EYNKASNDLYNFRVVDILKSQEQITD 79

```

```

RESULT 3
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OR INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-431-480-4

```

Query Match	53.8%	Score 77	DB 3	Length 142
Best Local Similarity	53.8%	Pred. No. 8	1e-05	
Matches	14	Conservative	6	Mismatches 6, Indels 0, Gaps 0
QY	2 QYNKESDDKHTFRFLRVAKNQKQVLD	27		
	: : : : : : : : : : : : : : : : :			
Db	55 EYNKESDDKVFVLVVKTLQAOQLQVTN	80		

```

RESULT 4
US-09-617-302-4
: Sequence 4, Application US/09617302
: Patent NO. 6245529
: GENERAL INFORMATION:
: APPLICANT: Holloway, James L.
: APPLICANT: Feldhaus, Andrew
: TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
: FILE REFERENCE: 98-72 C1
: CURRENT APPLICATION NUMBER: US/09/617,302
: CURRENT FILING DATE: 2000-07-17
: PRIOR APPLICATION NUMBER: 09/431,480
: PRIOR FILING DATE: 1999-11-01
: PRIOR APPLICATION NUMBER: 60/109,217
: PRIOR FILING DATE: 1998-11-20
: PRIOR APPLICATION NUMBER: 60/156,382
: PRIOR FILING DATE: 1999-09-28
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 142
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-617-302-4

```

Query Match	53.8%	Score 77;	DB 3;	Length 142;
Best Local Similarity	53.8%;	Pred. No. 8.1e-05;		
Matches 14; Conservative	6;	Mismatches 6;	Indels 0;	Gaps 0;

```

Qy      2 QYNKESDDKTHFRIFRLVLRVQRQVTD 27
      : ||||: ||| : : | |||:
Db      55 EYNKESDDKYVFLVVKTLQAQLQVTN 80

```

RESULT 5
US-09-431-480-3

```

: TITLE OF INVENTION: TESTS SPECIFIC CYSPATIN-LIKE PROTEIN CYSTATIN
: FILE REFERENCE: 98-72
: CURRENT APPLICATION NUMBER: US/09/431,480
: EARLIER FILING DATE: 1999-11-01
: EARLIER APPLICATION NUMBER: 60/109,217
: EARLIER FILING DATE: 1998-11-20
: EARLIER APPLICATION NUMBER: 60/156,382
: EARLIER FILING DATE: 1999-09-28
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 142
: TYPE: PRT
: ORGANISM: Mus musculus
: US-09-431-480-3

```

Query	2 QYNKESDDKTHFRIFRVLTNKQROQVTD	27
	: : : : : :	
DB	55 EYNKESDDKTVPLVDKTLHAKQITLD	80

```

RESULT 6
US-09-617-302-3
Sequence 3, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 142
TYPE: PRT
ORGANISM: Mus musculus
US-09-617-302-3

```

Query Match Similarity 53.1%; Score 76; DB 3; length 142;
Best Local Similarity 50.0%; Pred. No. 0.00012;
Matches 13; Conservative 7; Mismatches 6; Indels 0;
Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 14:07:11 ; Search time 6.91071 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-9
Perfect score: 143
Sequence: 1 DQYKESDDKXHFRLVLRKQOVTD 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/iaa/5A.COMB.dep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.dep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.dep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.dep.*
5: /cgn2_6/ptodata/2/iaa/PTCUS.COMB.dep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.dep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	56.6	141	3	US-09-431-480-2
2	81	56.6	141	3	US-09-617-302-2
3	77	53.8	142	3	US-09-431-480-4
4	77	53.8	142	3	US-09-617-302-4
5	76	53.1	142	3	US-09-431-480-3
6	76	53.1	142	3	US-09-617-302-3
7	66	46.2	116	4	US-09-775-932-16
8	66	46.2	121	4	US-09-775-932-8
9	66	46.2	139	2	US-08-791-522-4
10	66	46.2	139	3	US-09-314-777-4
11	66	46.2	139	4	US-08-849-303-15
12	66	46.2	141	3	US-08-744-138-6
13	66	46.2	141	4	US-09-241-376-6
14	66	46.2	141	4	US-09-940-497-6
15	66	46.2	141	4	US-08-849-303-24
16	65	45.5	120	4	US-09-775-932-2
17	65	45.5	120	6	5432264-4
18	65	45.5	145	2	US-08-832-535-11
19	65	45.5	146	2	US-08-791-522-3
20	65	45.5	146	3	US-08-744-138-3
21	65	45.5	146	3	US-09-019-485-4
22	65	45.5	146	3	US-09-314-777-3
23	65	45.5	146	3	US-09-431-480-6
24	65	45.5	146	3	US-09-617-302-6
25	65	45.5	146	3	US-09-241-376-3
26	65	45.5	146	4	US-09-528-436B-3
27	65	45.5	146	4	US-09-886-319A-47

28	65	45.5	146	4	US-09-940-497-3	Sequence 3, Appli
29	65	45.5	146	4	US-09-976-594-37	Sequence 37, Appli
30	65	45.5	146	4	US-08-849-303-17	Sequence 17, Appli
31	65	45.5	146	5	PCT-US95-07135-9	Sequence 9, Appli
32	65	45.5	146	6	5432264-6	Patent No. 5432264
33	63	44.1	56	1	US-08-538-165A-4	Sequence 4, Appli
34	62	43.4	112	4	US-08-849-303-16	Sequence 16, Appli
35	62	43.4	118	4	US-09-775-932-24	Sequence 24, Appli
36	62	43.4	127	4	US-08-849-303-19	Sequence 19, Appli
37	62	43.4	140	3	US-09-431-480-5	Sequence 5, Appli
38	62	43.4	140	3	US-09-617-302-5	Sequence 46, Appli
39	62	43.4	140	4	US-09-886-319A-46	Sequence 46, Appli
40	62	43.4	140	4	US-09-886-319A-48	Sequence 48, Appli
41	62	43.4	140	4	US-08-849-303-18	Sequence 18, Appli
42	61	42.7	111	4	US-09-775-932-18	Sequence 18, Appli
43	56	39.2	162	4	US-08-849-303-25	Sequence 25, Appli
44	54	37.8	56	1	US-08-538-165A-1	Sequence 1, Appli
45	54	37.8	56	1	US-08-538-165A-7	Sequence 7, Appli

ALIGNMENTS

```

RESULT 1
US-09-431-480-2
; Sequence 2, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-431-480-2

Query Match      56.6%; Score 81; DB 3; Length 141;
Best local Similarity 53.8%; Pred. No. 1.9e-05;
Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Cy      2 DQYKESDDKXHFRLVLRKQOVTD 27
Db      54 EYNKASNDLYNFRVVDLKSQEOITD 79

RESULT 2
US-09-617-302-2
; Sequence 2, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

QY 1 DOYNKESDDKXHFRIFRVLKVRQ 24
 |||
 Db 27 DOYNKESDDKXHFRIFRVLKVRQ 50

RESULT 15

AAMS4871
 ID AAMS4871 standard; protein; 50 AA.

AC AAMS4871;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26976.

KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-02344687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-483446/52.

XX DR

XX PT

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains.

XX Example 4; SEQ ID NO 26976; 650bp + Sequence Listing; English.

XX PS

XX The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC brain. They can be used to measure gene expression in brain cell samples,

XX CC which may enable the diagnosis and improved treatment of nervous system

XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX CC the probes of the invention

XX SQ Sequence 50 AA;

QY Query Match 89.5%; Score 128; DB 4; Length 50;

Best Local Similarity 100.0%; Pred. No. 2.9e-12;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DOYNKESDDKXHFRIFRVLKVRQ 24

27 DOYNKESDDKXHFRIFRVLKVRQ 50

Search completed: March 18, 2004, 14:14:59

Job time : 24.4821 secs

CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labeled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 50 AA;

Query Match 89.5%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2,9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKYHFRIFRVLKVRQ 24
DB 27 DQYNKESDDKYHFRIFRVLKVRQ 50

RESULT 13
ID ABB19524 standard; protein; 50 AA.
XX ABB19524;
AC ABB19524;
XX 23-JAN-2002 (first entry)
DT
XX Protein #1523 encoded by probe for measuring heart cell gene expression.
DE
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX Homo sapiens.
XX OS
XX PN MO200157274-A2.
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001MO-US000666.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI, 2001-488990/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX PT
XX Claim 15; SEQ ID NO 21294; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease, the
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 50 AA;

Query Match 89.5%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2,9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKYHFRIFRVLKVRQ 24
DB 27 DQYNKESDDKYHFRIFRVLKVRQ 50

RESULT 14
ID AAM67252 standard; protein; 50 AA.
XX AAM67252;
AC AAM67252;
XX 06-NOV-2001 (first entry)
DT
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.
DE
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS
XX PN WO200157276-A2.
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001MO-US000668.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI, 2001-488990/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX PT
XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.
XX PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX SQ Sequence 50 AA;

Query Match 89.5%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2,9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PT
PS Claim 27; SEQ ID NO 27814; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see A13315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 50 AA;
XX
Query Match 89.5%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKYHPRIFRVLKQVQRQ 24
Db 27 DQYNKESDDKYHPRIFRVLKQVQRQ 50
XX
RESULT 11
ABR32389
ID ABR32389 standard; peptide; 50 AA.
XX
AC ABR32389;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
KM cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
DR New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 15357; 327pp + Sequence listing; English.

XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;
XX
Query Match 89.5%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKYHPRIFRVLKQVQRQ 24
Db 27 DQYNKESDDKYHPRIFRVLKQVQRQ 50
XX
RESULT 12
ABR28913
ID ABR28913 standard; peptide; 50 AA.
XX
AC ABR28913;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
KM cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
DR New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 11881; 327pp + Sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon

RESULT 8	
AAM15096	
ID AAM15096 standard; protein; 50 AA.	
XX	
AC AAM15096;	
XX	
DT 12-OCT-2001 (first entry)	
DE Peptide #1530 encoded by probe for measuring cervical gene expression.	
XX	
KW Probe; human; microarray; gene expression; cervical epithelial cell;	
KW cervical cancer.	
XX	
OS Homo sapiens.	
XX	
WO200157278-A2.	
FN	
PD 09-AUG-2001.	
PD	
PF 30-JAN-2001; 2001WO-US000670.	
PR	
PR 04-FEB-2000; 2000US-0180312P.	
PR 26-MAY-2000; 2000US-0207456P.	
PR 30-JUN-2000; 2000US-00608408.	
PR 03-AUG-2000; 2000US-00632366.	
PR 21-SEP-2000; 2000US-0234687P.	
PR 27-SEP-2000; 2000US-0236359P.	
PR 04-OCT-2000; 2000GB-00024263.	
XX	
PA (MOLE-) MOLECULAR DYNAMICS INC.	
XX	
PI Penn SG, Hanzel DK, Chen W, Rank DR;	
XX	
WI; 2001-488901/53.	
XX	
PT Human genome-derived single exon nucleic acid probes useful for analyzing	
PT gene expression in human cervical epithelial cells.	
XX	
XS Claim 27; SEQ ID NO 19922; 487bp; English.	
XX	
CC The present invention relates to human single exon nucleic acid probes	
CC (SNP: see AAI10068-AAI28459). The present sequence is a peptide encoded	
CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs	
CC can be used to produce a single exon microarray, which can be used for	
CC measuring human gene expression in a sample derived from human cervical	
CC epithelial cells. By measuring gene expression, the probes are therefore	
CC useful in grading and/or staging of diseases of the cervix, notably	
CC cervical cancer. Note: The sequence data for this patent did not form	
CC part of the printed specification, but was obtained in electronic format	
CC directly from WIPO at fep.wipo.int/pub/published_pct_sequences	
SC Sequence 50 AA;	
Query Match 89.5%; Score 128; DB 4; Length 50;	
Best Local Similarity 100.0%; Pred. No. 2.9e-12;	
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 DQYNKESDDKXHFRIFFVLKXQRQ 24	
Db 27 DQYNKESDDKXHFRIFFVLKXQRQ 50	
RESULT 9	
ABB34086	
ID ABB34086 standard; peptide; 50 AA.	
XX	
AC ABB34086;	
XX	
DT 04-FEB-2002 (first entry)	
DE Peptide #1592 encoded by human foetal liver single exon probe.	
XX	
Human; foetal liver; gene expression; single exon nucleic acid probe.	

```

XX Homo sapiens.
OS
XX MO200157277-A2.
XX
XX PD 09-AUG-2001.
XX PE 30-JAN-2001; 2001WO-US000669.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human fetal liver. The
XX CC present sequence is a peptide encoded by a single exon nucleic acid probe
XX CC of the invention. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 50 AA;
SQ
Query Match 89.5%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKKHFRIFRVLKVQRQ 24
| | | | | | | | | | | | | | | | | |
DB 27 DQYNKESDDKKHFRIFRVLKVQRQ 50
RESULT 10
AAM27545
ID AAM27545 standard; protein; 50 AA.
XX AC AAM27545;
XX XX
XX JT 17-OCT-2001 (first entry)
XX DE Peptide #1582 encoded by probe for measuring placental gene expression.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
```

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYNKESDDKXHFRIFRVLKVRQVTD 27
 Db 32 DQYNKESDDKXHFRIFRVLKVRQVTD 58

RESULT 6

AAU79852

ID AAU79852 standard; protein; 137 AA.

AC AAU79852;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8).

KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation.

OS Homo sapiens.

PN WO200220567-A2.

PD 14-MAR-2002.

PF 29-AUG-2001; 2001MO-US026668.

PR 01-SEP-2000; 2000US-0230230P.

PA (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting

PT spermatogenesis, and inhibiting cancer procoagulant protein which leads

PS Claim 2; Page 93-94; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant

CC protein in an individual and thus inhibiting the thrombotic events

CC associated with cancer; promoting spermatogenesis, modulating seminal

CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm

CC motility and fertilisation; and as antigenic peptides to generate

CC antibodies. Zcys8 is useful as research reagent for characterising sites

CC of interaction between Zcys8 and its receptor. Zcys8 is useful in

CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

CC presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the

CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This is the amino acid sequence of human cystatin-8 (Zcys8)

CC Sequence 137 AA;

Query Match 100.0%; Score 143; DB 5; Length 137;

Best Local Similarity 100.0%; Pred. No. 4,4e-14;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DQYNKESDDKXHFRIFRVLKVRQVTD 27

Db 52 DQYNKESDDKXHFRIFRVLKVRQVTD 78

RESULT 7

AAU79865

ID AAU79865 standard; peptide; 80 AA.

AC AAU79865;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #13.

KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.

OS Homo sapiens.

PN WO200220567-A2.

PD 14-MAR-2002.

PF 29-AUG-2001; 2001MO-US026668.

PR 01-SEP-2000; 2000US-0230230P.

PA (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting

PT spermatogenesis, and inhibiting cancer procoagulant protein which leads

PS Claim 2; Page 98; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant

CC protein in an individual and thus inhibiting the thrombotic events

CC associated with cancer; promoting spermatogenesis, modulating seminal

CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm

CC motility and fertilisation; and as antigenic peptides to generate

CC antibodies. Zcys8 is useful as research reagent for characterising sites

CC of interaction between Zcys8 and its receptor. Zcys8 is useful in

CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

CC presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the

CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

CC Sequence 80 AA;

Query Match 95.8%; Score 137; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 2e-13;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DQYNKESDDKXHFRIFRVLKVRQVTD 27

Db 1 DQYNKESDDKXHFRIFRVLKVRQVTD 26

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 49 AA;

Query Match 100.0%; Score 143; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYNKESDDKXHFRIFRVLKVRQVTD 27
DB 1 DQYNKESDDKXHFRIFRVLKVRQVTD 27

RESULT 4

ID AAV79853 standard; protein; 115 AA.

AAV79853;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #1.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
sperm motility; fertilisation; antigenic fragment.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
spermatogenesis, and inhibiting cancer procoagulant protein which leads
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 94; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
protein in an individual and thus inhibiting the thrombotic events
associated with cancer; promoting spermatogenesis, modulating seminal
fluid viscosity, enhancing viability of cryopreserved sperm, sperm
motility and fertilisation; and as antigenic peptides to generate
antibodies. Zcys8 is useful as research reagent for characterising sites
of interaction between Zcys8 and its receptor. Zcys8 is useful in
enhancing fertilisation during assisted reproduction in humans and in
animals. Anti-(I) antibodies are useful to screen biological samples like
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
presence of Zcys8. The antibodies are also useful to isolate large
quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
The polynucleotide encoding (I) is useful to detect and to localise the
expression of a Zcys8 gene in a biological sample and Zcys8
oligonucleotide probes are useful for in vivo diagnosis. The
polynucleotide encoding (I) is useful in determining whether a subject's
chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
copy number changes, insertions, deletions, restriction site changes and
rearrangements and genetic alterations that inactivate the Zcys8 gene.
This sequence represents an antigenic fragment of human cystatin-8
(Zcys8)

Sequence 115 AA;

Query Match 100.0%; Score 143; DB 5; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.6e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYNKESDDKXHFRIFRVLKVRQVTD 27
DB 30 DQYNKESDDKXHFRIFRVLKVRQVTD 56

RESULT 5

ID AAV79854 standard; protein; 117 AA.

AAV79854;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #2.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
sperm motility; fertilisation; antigenic fragment.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
spermatogenesis, and inhibiting cancer procoagulant protein which leads
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 94-95; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
protein in an individual and thus inhibiting the thrombotic events
associated with cancer; promoting spermatogenesis, modulating seminal
fluid viscosity, enhancing viability of cryopreserved sperm, sperm
motility and fertilisation; and as antigenic peptides to generate
antibodies. Zcys8 is useful as research reagent for characterising sites
of interaction between Zcys8 and its receptor. Zcys8 is useful in
enhancing fertilisation during assisted reproduction in humans and in
animals. Anti-(I) antibodies are useful to screen biological samples like
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
presence of Zcys8. The antibodies are also useful to isolate large
quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
The polynucleotide encoding (I) is useful to detect and to localise the
expression of a Zcys8 gene in a biological sample and Zcys8
oligonucleotide probes are useful for in vivo diagnosis. The
polynucleotide encoding (I) is useful in determining whether a subject's
chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
copy number changes, insertions, deletions, restriction site changes and
rearrangements and genetic alterations that inactivate the Zcys8 gene.
This sequence represents an antigenic fragment of human cystatin-8
(Zcys8)

Sequence 117 AA;

Query Match 100.0%; Score 143; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 27 AA:

Query Match 100.0%; Score 143; DB 5; Length 27;
 Best Local Similarity 100.0%; Pred. No. 7.1e-15;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DOYNKESDDKXHFRIFRVLKVRQVTD 27
 |||||
 Db 1 DOYNKESDDKXHFRIFRVLKVRQVTD 27

RESULT 2

AAU79858
 ID AAU79858 standard; peptide; 35 AA.

XX AAU79858;

DT 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #6.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

OS WO200220567-A2.

PN 14-MAR-2002.

XX 29-AUG-2001; 2001MO-US026868.

PF 01-SEP-2000; 2000US-0230230P.

PR (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

PT WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 96; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 35 AA:

Query Match 100.0%; Score 143; DB 5; Length 35;
 Best Local Similarity 100.0%; Pred. No. 9.5e-15;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DOYNKESDDKXHFRIFRVLKVRQVTD 27
 |||||
 Db 9 DOYNKESDDKXHFRIFRVLKVRQVTD 35

RESULT 3

AAU79863
 ID AAU79863 standard; peptide; 49 AA.

XX AAU79863;

DT 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #11.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

OS WO200220567-A2.

PN 14-MAR-2002.

XX 29-AUG-2001; 2001MO-US026868.

PF 01-SEP-2000; 2000US-0230230P.

PR (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

PT WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 97-98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 14:02:15 ; Search time 24.4821 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-9

Perfect score: 143

Sequence: 1 DQYNKESDDKYHFRIFRVLKVRQVTD 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	27	5	AAU79859 Human cys
2	143	100.0	35	5	AAU79858 Human cys
3	143	100.0	49	5	AAU79863 Human cys
4	143	100.0	115	5	AAU79853 Human cys
5	143	100.0	117	5	AAU79854 Human cys
6	143	100.0	137	5	AAU79852 Human cys
7	137	95.8	80	5	AAU79865 Human cys
8	128	89.5	50	4	AAU79852 Human cys
9	128	89.5	50	4	AAU79852 Human cys
10	128	89.5	50	4	AAU79852 Human cys
11	128	89.5	50	4	AAU79852 Human cys
12	128	89.5	50	4	AAU79852 Human cys
13	128	89.5	50	4	AAU79852 Human cys
14	128	89.5	50	4	AAU79852 Human cys
15	128	89.5	50	4	AAU79852 Human cys
16	128	89.5	50	4	AAU79852 Human cys
17	128	89.5	50	4	AAU79852 Human cys
18	128	89.5	50	4	AAU79852 Human cys
19	128	89.5	50	4	AAU79852 Human cys
20	119	83.2	52	5	AAU79864 Human cys
21	81	56.6	141	3	AAV96576 Murine cy
22	81	56.6	141	4	AAE02403 Murine cy
23	81	56.6	141	4	AAE04432 Mouse tes
24	77	53.8	92	2	AAW78259 Fragment
25	77	53.8	113	6	ADA57563 Human sec

26	77	53.8	113	6	ADA41457 Human sec
27	77	53.8	113	7	ADC74577 Human sec
28	77	53.8	113	7	ADD38088 Human sec
29	77	53.8	114	2	AAW78153 Human sec
30	77	53.8	123	2	AAW78260 Fragment
31	77	53.8	142	2	AAW78258 Fragment
32	77	53.8	142	4	AAE02405 Human cys
33	77	53.8	142	4	AAE04434 Human cys
34	77	53.8	142	6	ADA57231 Human sec
35	77	53.8	142	6	ADA41112 Human sec
36	77	53.8	142	7	ADC74335 Human sec
37	77	53.8	142	7	ADD37980 Human sec
38	77	53.8	142	7	ADD46706 Human pro
39	77	53.8	142	7	ADD46710 Human pro
40	76	53.1	142	7	AAE02404 Murine cy
41	76	53.1	142	4	AAE04433 Mouse cys
42	76	53.1	143	6	ADA14374 Mouse spe
43	74	51.7	181	5	ABP41476 Human ova
44	67	46.9	116	3	AAE01203 Human mut
45	67	46.9	121	3	AAE01165 Human mut

ALIGNMENTS

RESULT 1
AAU79859 standard; peptide; 27 AA.

AAU79859;
15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #7.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
sperm motility; fertilisation; antigenic peptide.

Homo sapiens.
WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;
WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
spermatogenesis, and inhibiting cancer procoagulant protein which leads
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 96; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
protein in an individual and thus inhibiting the thrombotic events
associated with cancer; promoting spermatogenesis, modulating seminal
fluid viscosity, enhancing viability of cryopreserved sperm, sperm
motility and fertilisation; and as antigenic peptides to generate
antibodies. Zcys8 is useful as research reagent for characterizing sites
of interaction between Zcys8 and its receptor. Zcys8 is useful in
enhancing fertilisation during assisted reproduction in humans and in
animals. Anti-(I) antibodies are useful to screen biological samples like
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
presence of Zcys8. The antibodies are also useful to isolate large

THIS PAGE BLANK (USPTO)

Qy 7 ITDQYNKESDDKXHFRIFRVLKVGROV 33
 Db 68 VLTQVNOQSNDEYHLMPIKLVSSQV 94

RESULT 13

Q80125 PRELIMINARY; PRT; 464 AA.
 AC 080125;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Refun long form.
 OS Cyrtinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 NCBI_TaxId=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tsai P.-L., Chang G.-D., Huang C.-J.;
 RT "Purification and cloning of carp fetuin."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY225865; AA074862.1; -.
 DR GO; GO:0005874; C:mitochondrion; IEA.
 DR GO; GO:0004669; F:cysteine protease inhibitor activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0007018; P:microtubule-based movement; IEA.
 DR InterPro: IPR002453; Beta.tubulin.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00228; TUBULIN B AUTOREG; 1.
 SQ SEQUENCE 464 AA; 51698 MW; 7A54F71E4050895 CRC64;

Query Match 30.5%; Score 57; DB 13; Length 464;
 Best Local Similarity 32.3%; Pred. No. 14;
 Matches 10; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 2 DSLQWITDQYNKESDDKXHFRIFRVLKVGROV 32
 Db 158 ESVKTAIQKFNKESDDKXHFRIFRVLKVGRISTQ 188

RESULT 14

Q80389 PRELIMINARY; PRT; 96 AA.
 AC 080389;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cystein proteinase inhibitor.
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 NCBI_TaxId=3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamakawa S.;
 RT "Cysteine proteinase inhibitor."
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB014760; BAA28867.1; -.
 DR PIR: T07822; T07822.
 DR GO; GO:0004669; F:cysteine protease inhibitor activity; IEA.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00267; CYSTATIN; 1.
 SQ SEQUENCE 96 AA; 10973 MW; 027252E14BB6C4F4 CRC64;

Query Match 29.9%; Score 56; DB 10; Length 96;
 Best Local Similarity 33.3%; Pred. No. 3.2;

Matches 11; Conservative 6; Mismatches 14; Indels 2; Gaps 1;
 Qy 1 KDSLQWITDQYNKESDDKXHFRIFRVLKVGROV 33
 Db 21 KDIAEVAVAEYKNSQG--HHLTIVSILKCSQV 51

RESULT 15

Q9TY65 PRELIMINARY; PRT; 109 AA.
 AC 09TY65;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Antigen maltose binding protein (Fragment).
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 NCBI_TaxId=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94336252; PubMed=8058358;
 RA Trenholme K.R., Tree T.I., Gillespie A.J., Guderian R., Maizels R.M.,
 RA Bradley J.E.;
 RT "Heterogeneity of IgG antibody responses to cloned Onchocerca volvulus
 RT antigens in microfiladermia positive individuals from Bameraldas
 RT Province, Ecuador."
 RL Parasite Immunol. 16:201-209(1994).
 DR EMBL: S71364; AAC60509.1; -.
 DR GO; GO:0004669; F:cysteine protease inhibitor activity; IEA.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00267; CYSTATIN; 1.
 DR NON TER 1
 FT SEQUENCE 109 AA; 12701 MW; 3C6B5E9F14B082BB CRC64;

Query Match 29.9%; Score 56; DB 5; Length 109;
 Best Local Similarity 45.5%; Pred. No. 3.7;
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 12 NKESDDKXHFRIFRVLKVGROV 33
 Db 24 NEQSNDEYHLMPIKLVSSQV 45

Search completed: March 18, 2004, 14:21:27
 Job time : 21.9722 secs


```

FT CHAIN 21 140 CYSTATIN C.
FT VARIANT 16 16 A -> G.
FT VARIANT 84 84 L -> F.
SQ SEQUENCE 140 AA; 15517 MW; 3A563406DD58D785 CRC64;

Query Match
Best Local Similarity 35.8%; Score 67; DB 11; Length 140;
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

RESULT 10
Q9NH95 PRELIMINARY; PRT; 148 AA.
ID Q9NH95
AC Q9NH95;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Le-cystatin.
OS Litomosoides sigmodontis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Litomosoides.
OX NCBI_TaxID=42156;
RN [1]
RP SEQUENCE FROM N.A.
RA Pfeiff A.W., Hoffman W.H., Taylor D.W., Schulz-Key H.;
RT "Characterization and immunological properties of a cysteine protease
inhibitor of the filarial parasite litomosoides sigmodontis.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229173; AAF3896.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT CHAIN 25 148 LS-CYSTATIN.
SQ SEQUENCE 148 AA; 16686 MW; 2950AA89CA5339C9 CRC64;

Query Match
Best Local Similarity 32.6%; Score 61; DB 5; Length 148;
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKTHFRIFVLKVRQV 33
DB 51 QEMLPILTKYNGQSDAYHLMPIVLKVSQV 83

RESULT 11
Q812K6 PRELIMINARY; PRT; 4530 AA.
ID Q812K6
AC Q812K6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Pfl1500W.
OS plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Bertman M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Croun A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Feltwell T., Goble A., Goodhead I., Gilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornaby T., Holroyd S., Horrocks P.,
Humphrey S., Jagels K., James K.D., Johnson D., Kethrou A.,

```

```

RA Knights A., Kontorfov B., Kys S., Larie N., Lawson D., Lennart N.,
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ommond D., Price C., Quall M.A., Rabbittowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmidome falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:57-531(2002).
DR EMBL; AL829358; CAD51986.1; -
DR GO; GO:0009573; Cytibulose bisphosphate carboxylase complex; IEA.
DR GO; GO:0016984; F:ribulose-bisphosphate carboxylase activity; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0015977; P:carbon utilization by fixation of carbon di. . . ; IEA.
DR GO; GO:0006413; P:translational initiation; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000408; Reg_chir condense.
DR InterPro; IPR000685; Rubisco large.
DR InterPro; IPR006061; SBP dom1.
DR InterPro; IPR001950; TTF_SUI1.
DR PROSITE; PS0012; KC1_3; 1.
DR PROSITE; PS00157; RUBISCO_LARGE; 1.
DR PROSITE; PS00137; SBP_BACTERIAL_1; 1.
DR PROSITE; PS01118; SUI1_1; 1.
KW Hypothetical protein..
SQ SEQUENCE 4530 AA; 541990 MW; 372D9F6EB3317CBE CRC64;

Query Match 32.4%; Score 60.5; DB 5; Length 4530;
Best Local Similarity 38.2%; Pred. No. 55;
Matches 13; Conservative 9; Mismatches 9; Indels 3; Gaps 2;

Cy 1 KDSLGWITTDQTN-KESDDKTHFRFLRVLKXOROV 33
|::||| |:: ||:: ||:: ||::
Db 2113 KESIOWINDNPFNFISNDLS-LRLFKNKKIKINROI 2144

RESULT 12
ID 016159 PRELIMINARY; PRT; 161 AA.
AC 016159;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cystatin-type cysteine proteinase inhibitor.
GN BM-CPI-2.
OS Brugia malayi (filarial nematode worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchoercidae; Brugia.
OX NCBI_TaxID=6279;
[1]
RN RN SEQUENCE FROM N.A.
RP Gregory W.F., Blaxter M.L., Maizels R.M.;
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN RN SEQUENCE FROM N.A.
RP Gregory W.F., Maizels R.M.;
RA "Two distinct cystatin-type cysteine protease inhibitors from the
RT parasitic nematode Brugia malayi.";
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF015263; AAB69857.1; -.
DR EMBL; AF171713; AAD51086.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
SQ SEQUENCE 161 AA; 18406 MW; 8081351BBE226EB5 CRC64;

Query Match 30.5%; Score 57; DB 5; Length 161;
Best Local Similarity 37.0%; Pred. No. 4.1;
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
```

```

Db      47  ALWYAMKEYNKASNDLYNFRVVDILKSQEQTID  79

```

RESULT 6

ID	ORGANISM	PRELIMINARY	PRJ	112 AA
AC	096SR4			
AD	096SR4			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Cystatin (Fragment).			
OS	Acipenser sinensis (Chinese sturgeon).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;			
OC	Acipenser.			
OX	NCBI_TaxId=61970;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver:			
RA	Bai J., Lao H., Ye X., Li Y., Lou J.;			
RT	"Molecular cloning and sequence analysis of cystatin cDNA from two			
RT	species of sturgeons."			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF344610; AAK16731.1; -.			
DR	HSP; P01038; IAY0.			
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; CY; 1.			
DR	PROSITE; PS00287; CYSTATIN; 1.			
FT	NON TER			
FT	NON TER			
SEQUENCE	112 AA; 1231 MW; 48CEBFED9A08C00 CRC64;			

RESULT 7

ID	Q96SR3	PRELIMINARY;	PRT;	112 AA.
AC	O96SR3;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Cystatin (Fragment).			
OS	Acipenser schrenckii (Amur sturgeon).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Acipenseriformes; Acipenseridae;			
OC	Acipenser.			
NC	NCBI_TaxId=11304;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Bai J., Lao H., Ye X., Li Y., Lou J.;			
RT	"Molecular cloning and sequence analysis of cystatin cDNA from two			
RL	species of sturgeons.",			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	HMBL; AF34611; AAK16732.1; -.			
DR	HSSP; P01038; 1A90.			
DR	GO; GO:0004669; F:cysteine protease inhibitor activity; IEA.			
DR	Interpro: IPR000010; Cystatin.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	PROSITE; PS00287; CYSTATIN; 1.			
FT	NON TER			
FT	1			
FT	1			
SQ	SEQUENCE	112 AA;	12231 MW;	48ECBFEF8A08C0 CRC64;

Query Match 41.2%; Score 77; DB 13; Length 112;

Best Local Similarity 42.4%; Pred. No. 0.0042;
Matches 14; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

RESULT

ID	Q29212	PRELIMINARY;	PT;	81 AA.
AC	Q29212			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Cystatin C precursor (Fragment).			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxId=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Small intestine;			
RX	MEDLINE=96327607; Pubmed=8672129;			
RA	Winteroe A.K., Fredholm M., Davies W.;			
RT	"Evaluation and characterization of a porcine small intestine CDNA			
RT	library.";			
RL	Mamm. Genome 7:509-517(1996).			
DR	EMBL; F14794; CNA23262.1; --			
DR	HSSP; P01034; I996.			
DR	GO; GO:0004689; F:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00031; cystatin; 1.			
FT	NON_TER	81		
FT	SEQUENCE	81 AA; 8771 MW; 25583543411358A7 CRC64;		

RESULT 9

ID	09EPX9	PRELIMINARY;	PRT;	140 AA.
AC	09EPX9;			
DT	01-MAR-2001 (T-EMBLrel. 16, Created)			
DT	01-MAR-2001 (T-EMBLrel. 16, Last sequence update)			
DT	01-JUN-2003 (T-EMBLrel. 24, Last annotation update)			
DE	Cystatin C.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c;			
RX	MEDLINE=21010502; PubMed=1144350;			
RA	Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakansson K., Grubb A.,			
RA	Gage F.H.;			
RT	"NGF-2-Responsive neural stem cell proliferation requires CCG, a novel			
RT	autocrine/paracrine cofactor."			
RL	Neuron 28:385-397(2000).			
DR	EMBL; AF111741; AAG40283.1; -.			
DR	HSSP; P01034; 1G96.			
DR	GO; GO:0004699; P:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; C7; 1.			
DR	PROSITE; PS00287; CYSTATIN; 1.			

```

DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)
DE Sci3delta.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "Sci3: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF35481; AL7192.1; -;
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; Cystatin; 1.
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 92.0%; Score 172; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.5e-16;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKHFRIFRVLKVRQ 32
Db 45 KDSLOWITDQYNKESDDKHFRIFRVLKVRQ 76

RESULT 3
Q8K5A3 PRELIMINARY; PRT; 139 AA.
AC Q8K5A3;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)
DE Cystatin 11.
GN CST11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA Hamil K.G., Hall S.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF501290; AA021709.1; -;
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; Cystatin; 1.
DR SMART: SM00043; Cy; 1.
SQ SEQUENCE 139 AA; 16686 MW; E1E3DB786B4D08C CRC64;

Query Match 62.0%; Score 116; DB 11; Length 139;
Best Local Similarity 48.6%; Pred. No. 1.7e-08;
Matches 17; Conservative 16; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKHFRIFRVLKVRQVTD 35
Db 45 KETLEYTEBYNKSEDLNFRVLRIILIKEROMTN 79

RESULT 4
Q9DAP1 PRELIMINARY; PRT; 141 AA.
AC Q9DAP1;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)
DE 1700006C19Rik protein.
GN 1700006C19Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aoechi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Sakurai L.M., Stambli F., Suzuki R., Tomita M., Wagner L., Maehiro T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Maehima J., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseuki S.,
RA Hayaishiaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK005665; BAB24175.1; -;
DR HSSP: P01038; ICEW.
DR MGD; MGI:1916544; 1700006C19Rik.
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; Cystatin; 1.
DR SMART: SM00043; Cy; 1.
SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8BAC378C CRC64;

Query Match 44.4%; Score 83; DB 11; Length 141;
Best Local Similarity 45.5%; Pred. No. 0.00077;
Matches 15; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 3 SLOWITDQYNKESDDKHFRIFRVLKVRQVTD 35
Db 47 ALWYAMKEYNKASNDLNFRLVDILKSQEQITD 79

RESULT 5
Q80ZNS PRELIMINARY; PRT; 141 AA.
AC Q80ZNS;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, last annotation update)
DE RIKEN cDNA 1700006C19 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC048681; AA048681.1; -;
DR GO: GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; Cystatin; 1.
DR ProDom: PD001231; Cystatin_C/M; 1.
DR SMART: SM00043; Cy; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8BAC378C CRC64;

Query Match 44.4%; Score 83; DB 11; Length 141;
Best Local Similarity 45.5%; Pred. No. 0.00077;
Matches 15; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 3 SLOWITDQYNKESDDKHFRIFRVLKVRQVTD 35

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 20.9722 Seconds
(without alignments)
526.560 Million cell updates/sec

Title: US-09-941-314-8
Perfect score: 187
Sequence: 1 KDSIQWITDQYNKESDDKHFRIFRVLKVGQVTD 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187	100.0	138	4	Q8WXU6
2	172	92.0	103	4	Q8WXU5
3	116	62.0	139	11	Q8K5A3
4	83	44.4	141	11	Q9DAP1
5	83	44.4	141	11	Q80ZNS
6	77	41.2	112	13	Q98SRA
7	77	41.2	112	13	Q98SR3
8	69	36.9	81	6	Q29212
9	67	35.8	140	11	Q9EPX9
10	61	32.6	148	5	Q9NH95
11	60.5	32.4	4530	5	Q812K6
12	57	30.5	161	5	Q16159
13	57	30.5	464	13	Q801Z5
14	56	29.9	96	10	Q80389
15	56	29.9	109	5	Q9TY65
16	56	29.9	222	12	Q39107

17	54	28.9	222	12	Q39106	Q39106 bunyavirus
18	54	28.9	235	12	Q8JPR0	Q8JPR0 bunyavirus
19	54	28.9	400	16	Q9ZDT8	Q9ZDT8 rickettsia
20	54	28.9	518	10	Q22157	Q22157 arabidopsis
21	54	28.9	878	10	Q8RXP6	Q8RXP6 arabidopsis
22	54	28.9	1320	5	Q8ILJ3	Q8ILJ3 plasmodium
23	54	28.9	2646	12	Q9WPZ7	Q9WPZ7 gill-associ
24	53	28.3	132	2	Q86214	Q86214 exiguobacte
25	53	28.3	194	5	Q61819	Q61819 caenorhabdi
26	53	28.3	205	5	Q8T7Y7	Q8T7Y7 caenorhabdi
27	53	28.3	319	16	Q9CJ05	Q9CJ05 lactococcus
28	53	28.3	505	5	Q8IU00	Q8IU00 caenorhabdi
29	53	28.3	551	11	Q61192	Q61192 mus musculu
30	53	28.3	551	11	Q8VI97	Q8VI97 mus musculu
31	53	28.3	564	11	Q8C6B2	Q8C6B2 mus musculu
32	53	28.3	651	2	Q9RA18	Q9RA18 mycobacter
33	53	28.3	726	2	Q9LAE4	Q9LAE4 azospirillu
34	53	28.3	846	5	Q86S89	Q86S89 polypilaetro
35	53	28.3	3429	11	Q08614	Q08614 mus musculu
36	52	27.8	109	16	Q7TU55	Q7TU55 prochloroco
37	52	27.8	132	2	Q9F4C0	Q9F4C0 bacillus ce
38	52	27.8	132	2	Q70107	Q70107 bacillus ce
39	52	27.8	213	10	Q8LK68	Q8LK68 cucumis sat
40	52	27.8	272	5	Q9XW68	Q9XW68 caenorhabdi
41	52	27.8	279	16	Q8R5P1	Q8R5P1 fusobacteri
42	52	27.8	428	10	Q9FY09	Q9FY09 arabidopsis
43	52	27.8	450	16	Q921U9	Q921U9 rickettsia
44	52	27.8	1200	11	Q921B9	Q921B9 mus musculu
45	51.5	27.5	133	11	Q9D264	Q9D264 mus musculu

ALIGNMENTS

RESULT 1
ID Q8WXU6 PRELIMINARY; PRT; 138 AA.
AC Q8WXU6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE SC13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF35480; F: cytosine protease inhibitor activity; IEA.
DR GO: 0004689; F: cytosine protease inhibitor activity; IEA.
DR InterPro: IPR000010; Cysteatin.
DR Pfam: PF00031; Cysteatin; 1.
DR SMART: SM00043; Cy; 1.
SQ SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;
Query Match 100.0%; Score 187; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.6e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KDSIQWITDQYNKESDDKHFRIFRVLKVGQVTD 35
DB 45 KDSIQWITDQYNKESDDKHFRIFRVLKVGQVTD 79
RESULT 2
ID Q8WXU5 PRELIMINARY; PRT; 103 AA.
AC Q8WXU5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

Search completed: March 18, 2004, 14:16:10
Job time : 4.16667 secs

KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109
FT DISULFID 123 143

Query Match 36.9%; Score 69; DB 1; Length 146;
Best Local Similarity 35.5%; Pred. No. 0.016;
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKSDDKYHFRIFVLKVRQV 33
DB 52 ALDFAVGEYNKASNDMTHSRALQVVRKQI 82

RESULT 13
CYTC RAT
ID CYTC RAT STANDARD; PRT; 127 AA.
AC P14841;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Fragment).
GN C573.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Buffalo;
RX MEDLINE=90092122; PubMed=2689174;
RA Cole T., Dickson P.W., Ennard F., Averill F., Risbridger G.,
RA Gauthier F., Schreiber G.;
RT "The cDNA structure and expression analysis of the genes for the
RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
RT in rat brain."
RL Eur. J. Biochem. 186:35-42(1989).
RN [2]
RP SEQUENCE OF 8-127.
RX MEDLINE=90380276; PubMed=2400577;
RA Ennard F., Ennard A., Faucher D., Capony J.-P., Derancourt J.,
RA Brillard M., Gauthier F.;
RT "Rat cystatin C: the complete amino acid sequence reveals a site for
RT N-glycosylation."
RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
RN [3]
RP SEQUENCE OF 8-49.
RX MEDLINE=88313020; PubMed=3044831;
RA Ennard A., Ennard F., Faucher D., Gauthier F.;
RT "Two rat homologues of human cystatin C."
RL FEBS Lett. 236:475-478(1988).
RN [4]
RP SEQUENCE OF 8-20.
RC TISSUE=Sertoli cells;
RX MEDLINE=9225121; PubMed=1563513;
RA Ennard A., Ennard F., Guillou F., Gauthier F.;
RT "Production of the cysteine proteinase inhibitor cystatin C by rat
RT Sertoli cells."
RL FEBS Lett. 300:131-135(1992).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity. Known to inhibits cathepsin B,
CC H. and L.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L23572; AAB48011.1; --
CC HSSP; P01038; ICEW.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; cystatin; 1.
CC SMART; SM00043; CY; 1.
CC PROSITE; PS00287; CYSTATIN; 1.

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16957; CAA34831.1; --
CC DR PIR; S07085; S07085.
CC DR PIR; S10587; S10587.
CC DR HSSP; P01034; IG96.
CC DR InterPro; IPR000010; Cystatin.
CC DR Pfam; PF00031; cystatin; 1.
CC DR SMART; SM00043; CY; 1.
CC DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT NON TER 1 1
FT SIGNAL <1 7
FT CHAIN 8 127 CYSTATIN C.
FT ACT SITE 18 18 REACTIVE SITE.
FT SITE 62 66 SECONDARY AREA OF CONTACT.
FT DISULFID 80 90 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT CONFLICT 25 25 A -> E (IN REF. 2).
SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;

Query Match 35.8%; Score 67; DB 1; Length 127;
Best Local Similarity 35.5%; Pred. No. 0.026;
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKSDDKYHFRIFVLKVRQV 33
DB 33 ALDFAVSEYNGKSDAYHSGRAIQVVRKQL 63

RESULT 14
CYT CYPCA
ID CYT CYPCA STANDARD; PRT; 129 AA.
AC P35481;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin precursor (Ovarian cystatin) (P12).
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-57; 74-80 AND 92-129.
RX TISSUE=Ovary;
RC MEDLINE=96208938; PubMed=8829807;
RA Tsai Y.-J., Chang G.-D., Huang C.-J., Chang Y.-S., Huang F.-L.;
RT "Purification and molecular cloning of carp ovarian cystatin."
RL Comp. Biochem. Physiol. 113B:573-580(1996).
CC -1- FUNCTION: Cysteine proteinase inhibitor.
CC -1- TISSUE SPECIFICITY: Detected in ovary but not in all the other
CC examined tissues.
CC -1- PTM: THIS PROTEIN IS FURTHER PROCESSED BY AN INTRAMOLECULAR
CC CLEAVAGE TO PRODUCE TWO CHAINS LINKED BY A DISULFIDE BRIDGE.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L23572; AAB48011.1; --
CC HSSP; P01038; ICEW.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; cystatin; 1.
CC SMART; SM00043; CY; 1.
CC PROSITE; PS00287; CYSTATIN; 1.

[4]
RN. SEQUENCE FROM N.A.
RP.
RX MEDLINE=89350949; PubMed=2764935;
RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,
RA Isemura S., Sanada K.;
RT "The human cystatin C gene (CST3) is a member of the cystatin gene
RT family which is localized on chromosome 20.";
RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
RN 151


```

FT ACT_SITE 39 39 REACTIVE SITE.
FT SITE 83 87 SECONDARY AREA OF CONTACT.
FT DISULFID 101 111 BY SIMILARITY.
FT DISULFID 125 145 BY SIMILARITY.
SQ SEQUENCE 148 AA; 16346 MW; 1523C831169E5B9A CRC64;

Query Match 38.5%; Score 72; DB 1; Length 148;
Best Local Similarity 38.7%; Pred. No. 0.0064;
Matches 12; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRIFVLKVRQOV 33
Db 54 ALGFABSEYNKSNDRYHSRALQVVRAROI 84

RESULT 8
CYT_COTJA STANDARD; PRT; 116 AA.
AC P81061;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin (Egg-white cystatin).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE.
RC TISSUE=Egg white; PubMed=9276465;
RX MEDLINE=97420480; Mentele R., Eckerskorn C., Torquato R.,
RA Gerhartz B., Eng H.A., Machleidt W., Fritz H., Auerswald E.A.,
RA Wittman J., Kolb H.J., Machleidt W., Fritz H., Auerswald E.A.;
RT "Quail cystatin: isolation and characterization of a new member of
RT the cystatin family and its hypothetical interaction with cathepsin
RT B.";
RL FEBS Lett. 412:551-558(1997).
CC -!- FUNCTION: This protein binds tightly to and inhibits papain and
CC cathepsin B.
CC -!- SIMILARITY: Belongs to the cystatin family.
DR HSSP; P01038; ICEW.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Phosphorylation.
FT ACT_SITE 9 9 REACTIVE SITE.
FT SITE 53 57 SECONDARY AREA OF CONTACT.
FT DISULFID 71 81
FT DISULFID 95 115
FT MOD_RES 80 80 PHOSPHORYLATION.
SQ SEQUENCE 116 AA; 13093 MW; 48248621053A2P70 CRC64;

Query Match 38.0%; Score 71; DB 1; Length 116;
Best Local Similarity 38.7%; Pred. No. 0.0067;
Matches 12; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRIFVLKVRQOV 33
Db 24 ALQFAMAEYNKASNDKYSRVRIISAKQOL 54

RESULT 9
CYTC_MACMU STANDARD; PRT; 146 AA.
AC O19092;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CSF3.
OS Macaca mulatta (Rhesus macaque).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U51912; AAB64050.1; -.
DR HSSP; P01034; IG96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 93 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29DF26 CRC64;

Query Match 37.4%; Score 70; DB 1; Length 146;
Best Local Similarity 35.5%; Pred. No. 0.012;
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRIFVLKVRQOV 33
Db 52 ALDFAVSEYNKASNDMYHSRALQVVRARQI 82

RESULT 10
CYTC_SAISC STANDARD; PRT; 146 AA.
AC O19093;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CSF3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC

```

RL Biochem. J. 217:813-817(1984).
RN [3]
RP SEQUENCE OF 24-139.
RX MEDLINE=84110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
MA Machleidt W.;
RA "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
RT of cystatin from chicken egg white";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [4]
RP CHARACTERIZATION OF PROTEIN.
RX MEDLINE=83256421; PubMed=6409085;
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
MA Sunter D.C., Barrett A.J.;
RA "Cystatin, a protein inhibitor of cysteine proteinases. Improved
RT purification from egg white, characterization, and detection in
RL chicken serum";
RL Biochem. J. 211:129-138(1983).
RN [5]
RP DISULFIDE BONDS
RA Grubb A., Loeffberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
RT cystatin";
RL FEBS Lett. 170:370-374(1984).
RN [6]
RP PHOSPHORYLATION
RX MEDLINE=89252033; PubMed=2721673;
RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
MA Bode W.;
RA "The cysteine proteinase inhibitor chicken cystatin is a
RT phosphoprotein";
RL FEBS Lett. 248:162-168(1989).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=89052676; PubMed=3191914;
RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,
MA Brzin J., Kos J., Turk V.;
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
RT its possible mode of interaction with cysteine proteinases";
RL EMBO J. 7:2593-2599(1988).
RN [8]
RP STRUCTURE BY NMR.
RX MEDLINE=94087719; PubMed=8263912;
RA Dieckmann T., Mitechang L., Hofmann M., Kos J., Turk V.,
MA Auerwald E.A., Jeanicke R., Oschkinat H.;
RT "The structures of native phosphorylated chicken cystatin and of a
RL recombinant unphosphorylated variant in solution";
J. Mol. Biol. 234:1048-1059(1993).
CC -I- FUNCTION: This protein binds tightly to and inhibits a variety of
CC thiol proteases including ficin, papain, and cathepsins B, C, H,
CC and L. Although isolated from egg white, it is also present in
CC serum.
CC
CC -I- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05077; AAA48744.1; -;
DR PIR; A34456; UDCH.
DR PDB; 1CEW; 31-JAN-94.
DR PDB; 1A67; 27-MAY-98.
DR PDB; 1A90; 17-JUN-98.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.
FT SIGNAL 1 23

FT CHAIN 24 139 CYSTATIN.
FT ACT SITE 32 32 REACTIVE SITE.
FT SITE 76 80 SECONDARY AREA OF CONTACT.
FT DISULFID 94 104
FT DISULFID 118 138
FT MOD RES 103 103 PHOSPHORYLATION (PARTIAL).
FT STRAND 35 36
FT TURN 39 40
FT HELIX 42 51
FT TURN 52 52
FT HELIX 53 56
FT TURN 57 58
FT STRAND 63 77
FT STRAND 81 95
FT TURN 96 97
FT TURN 99 100
FT HELIX 101 108
FT STRAND 115 125
FT TURN 126 129
FT STRAND 130 139
SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;
Query Match 39.0%; Score 73; DB 1; Length 139;
Best Local Similarity 41.9%; Pred. NO. 0.0044;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
QY 3 SLOWTDQYNKSDDKYHFRFVLKVRQV 33
DB 47 ALQFMAEYNRASNDKYSRVVRVISAKRQL 77
RESULT 7
CYTC_RABIT STANDARD; PRT; 148 AA.
AC O97862;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CST3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Japanese white; TISSUE=Bone;
RX MEDLINE=98424349; PubMed=9753427;
RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,
RA Kawashima H.;
RT "Large scale isolation of osteoclast-specific genes by an improved
RT method involving the preparation of a subtracted cDNA library";
RL Genes Cells 3:459-475(1998).
CC -I- FUNCTION: This is a thiol proteinase inhibitor.
CC -I- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB009342; BAA75921.1; -;
DR HSSP; P01034; 1G96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; FALSE NEG.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 148 CYSTATIN C.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AL096677; CAC13170.1; -
 EMBL; AL096677; CAC17423.1; -
 HSP; P01038; 1A90.
 Genew; HGNC:15959; CST11.
 InterPro; IPR000010; Cystatin.
 Pfam; PF00031; cystatin; 1.
 SMART; SM00043; CY; 1.
 PROSITE; PS00287; CYSTATIN; FALSE NEG.
 Thiol protease inhibitor; Signal; Alternative splicing.
 SIGNAL 1 25 POTENTIAL.
 CHAIN 26 137 CYSTATIN 11.
 SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).
 DISULFID 93 101 BY SIMILARITY.
 DISULFID 114 134 BY SIMILARITY.
 CARBOHYD 131 131 N-LINKED (GLCNAC...) (POTENTIAL).
 VARSPLIC 76 110 Missing (in isoform 2).
 /FTId=VSP 001260.
 SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 187; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 9.3e-19;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKHFRIFRVLKVRQVTD 35
 DB 44 KDSLOWITDQYNKESDDKHFRIFRVLKVRQVTD 78

RESULT 2

CS11_MOUSE STANDARD; PRT; 139 AA.
 AC Q9D269;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin 11 precursor.
 CS11.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 Hayashizaki Y.
 RA Functional annotation of a full-length mouse cDNA collection.;
 RL Nature 409:685-690(2001).
 CC -/- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -/- SIMILARITY: Belongs to the cystatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AK020300; BAB32061.1; -
 HSP; P01034; 1G96.
 MGD; MGI:1925490; Cst11.
 InterPro; IPR000010; Cystatin.
 Pfam; PF00031; cystatin; 1.
 SMART; SM00043; CY; 1.
 PROSITE; PS00287; CYSTATIN; FALSE NEG.
 Thiol protease inhibitor; Signal.
 SIGNAL 1 28 POTENTIAL.
 CHAIN 29 139 CYSTATIN 11.
 SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).
 DISULFID 94 102 BY SIMILARITY.
 DISULFID 115 135 BY SIMILARITY.
 CARBOHYD 134 134 N-LINKED (GLCNAC...) (POTENTIAL).
 SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;

Query Match 64.2%; Score 120; DB 1; Length 139;
 Best Local Similarity 55.9%; Pred. No. 1.5e-09;
 Matches 19; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKHFRIFRVLKVRQVTD 34
 DB 45 KETLEYVTDQYNKESDDLYNFRILRLKIMKQVT 78

RESULT 3

CS18_HUMAN STANDARD; PRT; 142 AA.
 AC O60676;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin 8).
 CS18 OR CRSS.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=95344753; PubMed=7619504;
 RA Cornwall G.A., Hann S.R.;
 RT "Transient appearance of CRSS protein during spermatogenesis and caput epididymal sperm maturation.";
 RL Mol. Reprod. Dev. 41:37-46(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
 Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,
 Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 Ellington A.G., Frankland J.A., Fraser A., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 Grahame D.V., Griffiths C., Griffiths J., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 Hammond S., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 Lehesvaara M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 Marsh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 4.16667 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-8

Perfect score: 187

Sequence: 1 KDSLOWITQYNKESDDKYHFRIFRVLKQVQVTD 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187	100.0	137	CS11 HUMAN	Q9h12 homo sapien
2	120	64.2	139	CS11 MOUSE	Q9d269 mus musculus
3	82	43.9	142	CS18 HUMAN	O60676 homo sapien
4	77	41.2	142	CS18 MOUSE	P12766 mus musculus
5	74	39.6	141	CYT HUMAN	P09228 homo sapien
6	73	39.0	139	CYT CHICK	F01038 gallus galli
7	72	38.5	148	CYT RABIT	O97862 oryctolagus
8	71	38.0	116	CYT COTJA	P81061 coturnix co
9	70	37.4	146	CYT MACMU	O19092 macaca mula
10	70	37.4	146	CYT SAISC	O19093 sainiri sci
11	70	37.4	148	CYT BOVIN	P01035 bos taurus
12	69	36.9	146	CYT HUMAN	P01034 homo sapien
13	67	35.8	127	CYT RAT	P14841 rattus norv
14	67	35.8	129	CYT CYPCA	P35481 cyprinus ca
15	67	35.8	140	CYT MOUSE	P21460 mus musculus
16	64	34.2	142	CS18 RAT	O88969 rattus norv
17	61	32.6	141	CYT RAT	P19313 rattus norv
18	60	32.1	130	CYT ONCKE	Q98967 oncorhynch
19	60	32.1	130	CYT ONCMY	Q91195 oncorhynch
20	58	31.0	141	CYT HUMAN	P01037 homo sapien
21	58	31.0	141	CYT HUMAN	P01036 homo sapien
22	56	29.9	162	CYT ONCVO	P22085 onchocerca
23	56	29.9	165	CS1L HUMAN	Q9h14 homo sapien
24	54	28.9	235	NCAP BUNLC	P04873 bunyavirus
25	53	28.3	111	CYT BITAR	P08935 bitis ariet
26	53	28.3	132	MERR BACCE	P22853 bacillus ce
27	53	28.3	474	SYE VIPBA	Q87rl6 vibrio para
28	52.5	28.1	255	DA82 YEAST	P21705 saccharomyc
29	52.5	28.1	735	DHDM HTPSX	Q48303 hyphomicrob
30	52	27.8	504	GUNW ERWCA	Q59395 erwinia car
31	52	27.8	3433	UTRO HUMAN	P46939 homo sapien
32	51	27.3	474	SYE VIBVY	Q7mmw8 vibrio vuln
33	51	27.3	505	VP5_AHSV4	Q02168 african hor

RESULT 1

CS11_HUMAN STANDARD; PRT; 137 AA.

AC Q9H12; Q9H113;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cystatin 11 precursor.

GN CST11 OR CST8L

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor J.D.,

RA Levaeslao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramey H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,

RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.M., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray F.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

CC -1- SUBCELLULAR LOCATION: Secreted (Potential).

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=Q9H112-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q9H112-2; Sequence=VSP_001260;

CC Note=No experimental confirmation available;

CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

Search completed: March 18, 2004, 14:23:14
Job time : 6.73611 secs

A:Molecule type: protein
A:Residues: 29-38;76-80 <BIO>
A>Note: forms RSC-1, RSC-2 and RSC-3 with differing amino-terminals were found; a form B
C:Genetics:
A:Gene: Cyss
A:Introns: 76/3; 114/3
C:Superfamily: cystatin, cystatin homology
C:Keywords: cysteine proteinase inhibitor
F:1-28/Domain: signal sequence #status predicted <SIG>
F:28-141/Product: cystatin S, form RSC-3 #status experimental <MAT>
F:29-141/Product: cystatin S, form RSC-3 #status predicted <MAT>
F:30-141/Domain: cystatin homology <CYS>
F:31-141/Product: cystatin S, form RSC-2 #status predicted <MAT>
F:32-141/Product: cystatin S, form RSC-1 #status predicted <MAT>
F:94-104,118-138/Disulfide bonds: #status experimental

Query Match 32.4%; Score 61; DB 2; Length 141;
Best Local Similarity 34.4%; Pred. No. 0.58;
Matches 11; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 2 DSLQWITDQYNKESDDKYHFRIFRVLKQVQV 33
DB 46 EALYAVNEYNKESDLYLSRVVEKDVQKQV 77

RESULT 11
JC2040
cystatin - chum salmon
N:Alternate names: cysteine proteinase inhibitor
C:Species: Oncorhynchus keta (Chum salmon)
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 16-Jul-1999
C:Accession: JC2040
R:Koida, Y.; Noso, T.
BioSci. Biotechnol. Biochem. 58, 164-169, 1994
A:Title: The complete amino acid sequence of pituitary cystatin from chum salmon.
A:Reference number: JC2040; MUID:94162738; PMID:7764512
A:Accession: JC2040
A:Molecule type: protein
A:Residues: 1-111 <KOI>
C:Comment: The intracellular role of this protein is the inhibition of intralysosomal pH
C:Superfamily: cystatin; cystatin homology
C:Keywords: cysteine proteinase inhibitor
F:2-111/Domain: cystatin homology <CYS>
F:48-52/Region: inhibitory
F:89-109/Disulfide bonds: #status experimental

Query Match 32.1%; Score 60; DB 1; Length 111;
Best Local Similarity 30.3%; Pred. No. 0.61;
Matches 10; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKYHFRIFRVLKQVQV 33
DB 17 RDALQPAVVEHNKKTDMFVRQVAKVNVNAQKQV 49

RESULT 12
JC4918
cystatin precursor - chum salmon
C:Species: Oncorhynchus keta (Chum salmon)
C>Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jan-2000
C:Accession: JC4918
R:Yamashita, M.; Konagaya, S.
J. Biochem. 120, 483-487, 1996
A:Title: Molecular cloning and gene expression of chum salmon cystatin.
A:Reference number: JC4918; MUID:97058289; PMID:8902609
A:Accession: JC4918
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-132 <YAM>
A:Cross-references: DDBJ:DB6628
A:Experimental source: liver
C:Comment: This protein is a homolog of mammalian cystatin C, and acts as an extracellular
C:Superfamily: cystatin; cystatin homology

C:Keywords: liver
F:1-21/Domain: signal sequence #status predicted <SIG>
F:23-132/Domain: Cystatin homology <CYS>

Query Match 32.1%; Score 60; DB 2; Length 132;
Best Local Similarity 30.3%; Pred. No. 0.74;
Matches 10; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 KDSLQWITDQYNKESDDKYHFRIFRVLKQVQV 33
DB 38 RDALQPAVVEHNKKTDMFVRQVAKVNVNAQKQV 70

RESULT 13
E64005
hypothetical protein HI0284 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999
C:Accession: E64005
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: E64005
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-139 <TIGR>
A:Cross-references: GB:U32714; GB:L42023; NID:g1573241; PID:g1573253; TIGR:HI0284
C:Genetics:
A:Start codon: GTG
C:Superfamily: Haemophilus influenzae hypothetical protein HI0284

Query Match 31.6%; Score 59; DB 2; Length 139;
Best Local Similarity 45.0%; Pred. No. 1.1;
Matches 9; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 15 SDDKYHFRIFRVLKQVQV 34
DB 91 NDKKFFKIFSIKIQKLT 110

RESULT 14
UDHUP1
cystatin S precursor - human
N:Alternate names: cystatin SA-III; salivary acidic protein-1
C:Species: Homo sapiens (man)
C>Date: 25-Feb-1985 #sequence_revision 08-Feb-1996 #text_change 16-Jul-1999
C:Accession: S17667; S16500; A01272; A29603; S19280; A56608
R:Bobek, L.A.; Aguirre, A.; Levine, M.J.
Biochem. J. 278, 627-635, 1991
A:Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ and
A:Reference number: S17667; MUID:91378918; PMID:1898352
A:Accession: S17667
A:Molecule type: mRNA
A:Residues: 1-141 <BOB>
A:Cross-references: EMBL:X54667; NID:g30365; PIDN:CAA38478.1; PID:g30366
R:Lamkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.
Arch. Biochem. Biophys. 288, 664-670, 1991
A:Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle
A:Reference number: S16500; MUID:91378515; PMID:1898055
A:Accession: S16500
A>Status: preliminary
A:Molecule type: protein
A:Residues: 21-134, 'D', 136-141 <IHU>
R:Isemura, S.; Saitoh, E.; Sanada, K.
J. Biochem. 96, 489-498, 1984
A:Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal
A:Reference number: A91985; MUID:85054716; PMID:6501254
A:Accession: A01272
A:Molecule type: protein

A;Cross-references: EMBL:M59470; NID:g192911; PIDN:AAA63298.1; PID:g192912
C;Superfamily: cystatin; cystatin homology
F;29-140/Domain: cystatin homology <Cys>
F;93-103,117-137/Disulfide bonds: #status predicted

Query Match 35.8%; Score 67; DB 2; Length 140;
Best Local Similarity 35.5%; Pred. No. 0.091;
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRIFRVLKVQRQV 33
:
Db 46 ALDFAVSEYNGKSNDAYHSRAIQVVRRARKQL 76

RESULT 9
S62326
cystatin C, parotid glands - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S62326
R;Cimerman, N.; Drobnic Kosorok, M.; Korant, B.D.; Turk, B.; Turk, V.
Biol. Chem. Hoppe-Seyler 377, 19-23, 1996
A;Title: Characterization of cystatin C from bovine parotid glands: cysteine proteinase
A;Reference number: S62326; MUID:97083236; PMID:8929810
A;Accession: S62326
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-47 <CIN>
C;Superfamily: cystatin; cystatin homology

Query Match 35.3%; Score 66; DB 2; Length 47;
Best Local Similarity 31.2%; Pred. No. 0.038;
Matches 10; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 KDSLWITDVYNKESDDKYHFRIFRVLKVQRQ 32
:
Db 16 QEALSPAVSEFNKSNDAIYQSRVVVRARQK 47

RESULT 10
JQ1470
Cystatin S precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 16-Jul-1999
C;Accession: JQ1470; A31891; S05252; S15132
R;Cox, J.L.; Shaw, P.A.
Gene 110, 175-180, 1992
A;Title: Structure, organization and regulation of a rat cysteine proteinase inhibitor-er
A;Reference number: JQ1470; MUID:92165056; PMID:1537554
A;Accession: JQ1470
A;Molecule type: DNA
A;Residues: 1-141 <COX>
A;Cross-references: GB:W75281; NID:g294537; PIDN:AAA41068.1; PID:g294538
R;Shaw, P.A.; Cox, J.L.; Barka, T.; Naito, Y.
J. Biol. Chem. 263, 18133-18137, 1988
A;Title: Cloning and sequencing of cDNA encoding a rat salivary cysteine proteinase inhib
A;Reference number: A31891; MUID:89053983; PMID:3263967
A;Accession: A31891
A;Molecule type: mRNA
A;Residues: 1-ISTDLVLYISFEHETLSC, 27-141 <SHA>
R;Bedi, G.S.
Arch. Biochem. Biophys. 273, 245-253, 1989
A;Title: Amino acid sequence of an inducible cysteine proteinase inhibitor (cystatin) fr
A;Reference number: S05252; MUID:89334379; PMID:2757396
A;Accession: S05252
A;Molecule type: protein
A;Residues: 28-113,'QE',116-141 <BED>
A;Note: 87-Glu and 88-His were also found; disulfide bonds determined
R;Nishihara, T.; Ishibashi, K.; Abe, K.
Biochim. Biophys. Acta 1077, 346-354, 1991
A;Title: Isolation of three forms of cystatin from submandibular saliva of isoproterenol-
A;Reference number: S15132; MUID:91230145; PMID:1903068
A;Accession: S15132

A;Accession: A32732
A;Molecule type: protein
A;Residues: 27-76

R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst
A;Reference number: A60552; PMID:90193615; PMID:2315647
A;Accession: A60552
A;Molecule type: protein
A;Residues: 27-49, 'XX', 52-64 <OLA>
A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
e defective gene is not present in CSF but is found instead in amyloid deposits
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
A;Title: Different forms of human cystatin C.
A;Reference number: S10607; PMID:91025625; PMID:2222856
A;Accession: S10607
A;Molecule type: protein
A;Residues: 27-53 <POP>
A;Experimental source: urine, kidney disease
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als
R;Grubb, A.; Lofberg, H.; Barrett, A.J.
FEBS Lett. 170, 370-374, 1984
A;Title: The disulfide bridges of human cystatin C (gamma-trace) and chicken cystatin.
A;Reference number: S01462
A;Contents: annotation; disulfide bonds
R;Berti, P.J.; Storer, A.C.
Biochem. J. 302, 411-416, 1994
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility
A;Reference number: S55305; PMID:94379969; PMID:8092991
A;Accession: S55305
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-49, 106-146 <BER>
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl
I patients with certain autoimmune diseases.
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importar
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hermorri
C;Genetics:
A;Gene: GDB:CST3
A;Cross-references: GDB:119817; OMIM:105150
A;Map position: 20p11.2-20p11.2
A;Introns: 81/3; 119/3
C;Superfamily: cystatin; cystatin homology
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline
F;1-26/Domain: signal sequence #status predicted <SIG>
F;1-26/Product: cystatin C #status experimental <MAT>
F;35-146/Domain: cystatin homology <CVS>
F;35-146/Region: inhibitory #status predicted
F;29/Modified site: hydroxyproline (Pro) (partial) #status experimental
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 36.9%; Score 69; DB 1; Length 146;
Best Local Similarity 35.5%; Pred. No. 0.051;
Matches 11; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

Qy 3 SLOWITQYNKESDDKYHFRIFRVLKVRQV 33
: : : : ||| | : ||| : ||| : ||| :
Db 52 ALDPFAVEYKNASNDMYHSRALQVVRRKKI 82

RESULT 6
S10587
Cystatin C - rat
C;Species: Rattus sp. (rat)
C;Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C;Accession: S10587
R;Ennard, F.; Ennard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gauthi
Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990
A;Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat
A;Reference number: S10587; PMID:90380276; PMID:2400577
A;Accession: S10587
A;Status: preliminary

THIS PAGE BLANK (USPTO)

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDOYNKESDDK 18
 Db 19 KDSLOWITDOYNKESDDK 36

RESULT 14

US-10-264-049-2608
 ; Sequence 2608, Application US/10264049
 ; Publication No. US20040005579A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Birse et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P133P1
 ; CURRENT APPLICATION NUMBER: US/10/264,049
 ; PRIOR FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569
 ; PRIOR FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: US 60/209,467
 ; PRIOR FILING DATE: 2000-06-07
 ; NUMBER OF SEQ ID NOS: 4360
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 2608
 ; LENGTH: 181
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-264-049-2608

Query Match 42.2%; Score 79; DB 15; Length 181;
 Best Local Similarity 42.4%; Pred. No. 0.0051;
 Matches 14; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 KDSLOWITDOYNKESDDKXHFIRFVLKVRQV 33
 Db 85 QOALNEALSEYNKASNDAYHSRAMRVVRKQV 117

RESULT 15

US-09-775-932-8
 ; Sequence 8, Application US/09775932
 ; Patent No. US20020137671A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of British Columbia
 ; TITLE OF INVENTION: Production and use of Modified Cystatins
 ; FILE REFERENCE: 58069
 ; CURRENT APPLICATION NUMBER: US/09/775,932
 ; CURRENT FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: CA99/00717
 ; PRIOR FILING DATE: 1998-08-05
 ; PRIOR APPLICATION NUMBER: 60/095,503
 ; PRIOR FILING DATE: 1998-08-05
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 121
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-775-932-8

Query Match 39.6%; Score 74; DB 9; Length 121;
 Best Local Similarity 35.5%; Pred. No. 0.016;
 Matches 11; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 3 SLOWITDOYNKESDDKXHFIRFVLKVRQV 33
 Db 27 ALHFVISEYNKATEDEYRRLRLVLRREQI 57

Search completed: March 18, 2004, 14:30:53
 Job time : 21.1806 secs

```
Db 1 DQNKESDDKYHFRIFRVLKVRQVTD 27

RESULT 9
US-09-941-314-15
; Sequence 15, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-15

Query Match 73.3%; Score 137; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 QYNKESDDKYHFRIFRVLKVRQVTD 35
Db 1 QYNKESDDKYHFRIFRVLKVRQVTD 26

RESULT 10
US-09-941-314-10
; Sequence 10, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-10

Query Match 66.8%; Score 125; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NKESDDKYHFRIFRVLKVRQVTD 35
Db 1 NKESDDKYHFRIFRVLKVRQVTD 24

RESULT 11
US-09-941-314-14
; Sequence 14, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-14

Query Match 63.6%; Score 119; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KESDDKYHFRIFRVLKVRQVTD 35
Db 1 KESDDKYHFRIFRVLKVRQVTD 23

RESULT 12
US-09-941-314-7
; Sequence 7, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-7

Query Match 52.9%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.5e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDK 18
Db 1 KDSLOWITDQYNKESDDK 18

RESULT 13
US-09-941-314-6
; Sequence 6, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-6

Query Match 52.9%; Score 99; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
```

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
 OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUE 5.00e-23
 OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01
 US-09-864-761-34822

Query Match 92.0%; Score 172; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.9e-16;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKQVRQ 32
 |||||
 Db 19 KDSLOWITDQYNKESDDKYHFRIFRVLKQVRQ 50

RESULT 6
 US-09-864-761-48936
 ; Sequence 48936, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeonica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 48936
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

FEATURE:
 OTHER INFORMATION: MAP TO AL096677.18
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
 OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUE 5.00e-23
 OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01
 US-09-864-761-48936

Query Match 92.0%; Score 172; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.9e-16;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKQVRQ 32
 |||||
 Db 19 KDSLOWITDQYNKESDDKYHFRIFRVLKQVRQ 50

RESULT 7
 US-09-941-314-9
 ; Sequence 9, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR APPLICATION NUMBER: 60/230,230
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 27
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-9

Query Match 76.5%; Score 143; DB 9; Length 27;
 Best Local Similarity 100.0%; Pred. No. 9.8e-13;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DQYNKESDDKYHFRIFRVLKQVRQVTD 35
 |||||
 Db 1 DQYNKESDDKYHFRIFRVLKQVRQVTD 27

RESULT 8
 US-09-941-314-13
 ; Sequence 13, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR APPLICATION NUMBER: 60/230,230
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-13

Query Match 76.5%; Score 143; DB 9; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.9e-12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DQYNKESDDKYHFRIFRVLKQVRQVTD 35
 |||||

FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-3

Query Match 100.0%; Score 187; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 22 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 56

RESULT 3

US-09-941-314-4
Sequence 4, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-4

Query Match 100.0%; Score 187; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 4e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 24 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 58

RESULT 4

US-09-941-314-2
Sequence 2, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 137
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-2

Query Match 100.0%; Score 187; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.7e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 44 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 78

RESULT 5

US-09-864-761-34822
Sequence 34822, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34822
LENGTH: 50
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL109954.10
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 21.1806 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-8
Perfect score: 187
Sequence: 1 KDSLOWITDOYNKESDDKHFRIFRVLKVRQVTD 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187	100.0	35	9 US-09-941-314-8	Sequence 8, Appli
2	187	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
3	187	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
4	187	100.0	137	9 US-09-941-314-2	Sequence 2, Appli
5	172	92.0	50	9 US-09-864-761-34822	Sequence 34822, A
6	172	92.0	50	9 US-09-864-761-48936	Sequence 48936, A
7	143	76.5	27	9 US-09-941-314-9	Sequence 9, Appli
8	143	76.5	49	9 US-09-941-314-13	Sequence 13, Appli
9	137	73.3	80	9 US-09-941-314-15	Sequence 15, Appli
10	125	66.8	46	9 US-09-941-314-10	Sequence 10, Appli
11	119	63.6	52	9 US-09-941-314-14	Sequence 14, Appli
12	99	52.9	18	9 US-09-941-314-7	Sequence 7, Appli
13	99	52.9	36	9 US-09-941-314-6	Sequence 6, Appli
14	79	42.2	181	15 US-10-264-049-2608	Sequence 2608, Ap
15	74	39.6	121	9 US-09-775-932-8	Sequence 8, Appli

Sequence 24, Appli
Sequence 6, Appli
Sequence 16, Appli
Sequence 15, Appli
Sequence 4, Appli
Sequence 16, Appli
Sequence 24, Appli
Sequence 2, Appli
Sequence 17, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 47, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 18, Appli
Sequence 13, Appli
Sequence 18, Appli
Sequence 46, Appli
Sequence 48, Appli
Sequence 14, Appli
Sequence 23, Appli
Sequence 20, Appli
Sequence 22, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 21, Appli
Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-941-314-8
; Sequence 8, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-8

Query Match 100.0%; Score 187; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDSLOWITDOYNKESDDKHFRIFRVLKVRQVTD 35
Db 1 KDSLOWITDOYNKESDDKHFRIFRVLKVRQVTD 35

RESULT 2
US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

THIS PAGE BLANK (USPTO)

```
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
; US-08-791-522-4

Query Match 39.0%; Score 73; DB 2; Length 139;
Best Local Similarity 41.9%; Pred. No. 0.0027;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRVLKVRQV 33
Db 47 ALQFAMAENRASNDKYSSRVVRVISAKROL 77

RESULT 14
US-09-314-777-4
; Sequence 4, Application US/09314777
; Patent No. 6110686
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,777
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,522
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
; US-09-314-777-4

Query Match 39.0%; Score 73; DB 3; Length 139;
Best Local Similarity 41.9%; Pred. No. 0.0027;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRVLKVRQV 33
Db 47 ALQFAMAENRASNDKYSSRVVRVISAKROL 77

RESULT 15
US-08-849-303-15
; Sequence 15, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urwin, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-849-303-15

Query Match 39.0%; Score 73; DB 4; Length 139;
Best Local Similarity 41.9%; Pred. No. 0.0027;
Matches 13; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRVLKVRQV 33
Db 47 ALQFAMAENRASNDKYSSRVVRVISAKROL 77

Search completed: March 18, 2004, 14:25:36
Job time : 9.95833 secs
```


Query Match 39.6%; Score 74; DB 4; Length 141;
Best Local Similarity 35.5%; Pred. No. 0.002;
Matches 11: Conservative 12; Mismatches 8; Indels

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-2

Query Match      44.4%; Score 83; DB 3; Length 141;
Best Local Similarity 45.5%; Pred. No. 9.4e-05;
Matches 15; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

QY 3 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 47 ALWYAMKEYNKASNDLYNFRVVDILKSQEQITD 79

RESULT 3
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match      43.9%; Score 82; DB 3; Length 142;
Best Local Similarity 45.7%; Pred. No. 0.00013;
Matches 16; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQALQVTN 80

RESULT 4
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-4

Query Match      41.2%; Score 77; DB 3; Length 142;
Best Local Similarity 40.0%; Pred. No. 0.00072;
Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQALQVTN 80

RESULT 5
US-09-431-480-3
; Sequence 3, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-480-3

Query Match      41.2%; Score 77; DB 3; Length 142;
Best Local Similarity 40.0%; Pred. No. 0.00072;
Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQALQVTN 80

RESULT 6
US-09-617-302-3
; Sequence 3, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-617-302-3

Query Match      41.2%; Score 77; DB 3; Length 142;
Best Local Similarity 40.0%; Pred. No. 0.00072;
Matches 14; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
Db 46 KQCVWFAMQEYNKESDDKYVFLVVKTLQALQVTN 80
```

; ERROR FILING DATE: 1999-09-28
 ; NUMBER OF SEQ ID NOS: 22

Query Match	29.1%;	Score 55;	DB 10;	Length 665;
Best Local Similarity	33.3%;	Pred. No. 89;		

```

DR EMBL; AE002178; AAF38049.1; -.
DR EMBL; AF002547; BAA98780.1; -.
DR EMBL; AE017159; BAA98525.1; -.
DR PIR; B86562; B86562.
DR PIR; D72061; D72061.
DR TIGR; CF0175; -.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006025; Pept M.Zn.BS.
DR InterPro; IPR006025; Pept M.Zn.BS.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 184 AA; 22024 MW; EF36820DIA97A814 CRC64;

Query Match 30.4%; Score 57.5; DB 16; Length 184;
Best Local Similarity 38.5%; Pred. No. 10;
Matches 15; Conservative 6; Mismatches 11; Indels 7; Gaps 1;

QY 1 RKKTFLSVHEVMAVENYAKD-----SLQWITDQYNKE 32
Db 26 RKKTFLSHRVLARPPVDYCPGKIYDLOEIYEELNAQ 64

RESULT 9
O62173 PRELIMINARY; PRT; 209 AA.
AC O62173
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F15D3.6 protein.
GN F15D3.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Rhabditidae.
OC Rhabditidae; Paloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RA White S.;
RP SEQUENCE FROM N.A.
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81063; CAB02955.1; -.
DR PIR; T20975; T20975.
DR WormPep; F15D3.6; CE15853.
DR InterPro; IPR006797; MSF1.
DR Pfam; PF04707; MSF1; 1.
DR PROSITE; PS50904; PRELI_MSF1; 1.
SQ SEQUENCE 209 AA; 23667 MW; 346AB71D4BBD39C1 CRC64;

Query Match 29.6%; Score 56; DB 5; Length 209;
Best Local Similarity 35.5%; Pred. No. 18;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY 2 KKTFLSVHEVMAVENYAKDSLQWITDQYNKE 32
Db 140 EKTFLSIYSQNA--NKGQGVWVIDHLKKE 168

RESULT 10
Q81J78 PRELIMINARY; PRT; 175 AA.
AC Q81J78
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA polymerase sigma-70 factor, ECF subfamily.
GN BA5610.

OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T., Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R., Holtzaple E.K., Okstad O.A., Helgason E., Ristone J., Wu M., Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M., DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H., Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Redune D., Benson J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman W.C., Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C., Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B., Fraser C.M.;
RA "The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017041; AAP29248.1; -.
DR TIGR; BA5610; -.
DR InterPro; IPR009043; RNA_pol_sigma.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007630; Sigma70_r4.
DR Pfam; PF04542; sigma70_r2; 1.
DR Pfam; PF04545; sigma70_r4; 1.
KW Complete proteome.
SQ SEQUENCE 175 AA; 21395 MW; 58583126E6657D80 CRC64;

Query Match 29.1%; Score 55; DB 16; Length 175;
Best Local Similarity 39.5%; Pred. No. 20;
Matches 15; Conservative 4; Mismatches 7; Indels 12; Gaps 2;

QY 3 KTFLSVHEVMAVENYAKDSLQWIT-----DQYNKE 32
Db 40 KVFSSRHQLSVENYKK---WTTICVTFYDFYSKK 73

RESULT 11
O31093 PRELIMINARY; PRT; 327 AA.
AC O31093
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PSSG.
GN PSSG.
OS Rhizobium leguminosarum (biovar viciae).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=387;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF39;
RA Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VF39;
RX MEDLINE=99113394; PubMed=9914965;
RA Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G., Ksenzenko V.N.;
RT "Structural and functional organization of the exopolysaccharide biosynthesis genes in Rhizobium leguminosarum bv. viciae VF39.";
RL Mol. Biol. (Mosk) 32:797-804(1998).
DR EMBL; AF028810; AAB88894.1; -.
DR InterPro; IPR001173; Glyco_transf.
DR Pfam; PF00535; Glycos_transf_2; 1.
SQ SEQUENCE 327 AA; 36361 MW; E3AE2231DCF05A92 CRC64;

Query Match 29.1%; Score 55; DB 2; Length 327;
Best Local Similarity 47.8%; Pred. No. 41;

```

DR	SMART; SMO0421; HTH_LUXR; 1.
DR	PROSITE; PS00622; HTH_LUXR_FAMILY; 1.
KW	Complete proteome.
SQ	SEQUENCE 240 AA; 28117 MW; 1434A0B458CC54B0 CRC64;
Query Match	30.7%; Score 59; DB 16; Length 240;
Best Local Similarity	34.5%; Pred.No.12;
Matches	10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
OY	3 KTFLSVHEVMVAENVAKDSLQWITDQYNK 31 ::: ::: ::: ::: ::: ::: ::: ::: :::
Db	74 KNFLAIDPVLNPFNSQGHLMNDLDFNE 102 ::: ::: ::: ::: ::: ::: ::: ::: :::
RESULT 6	
Q81225	PRELIMINARY; PRT; 851 AA.
ID Q81225	
AC Q81225;	
DT 01-MAR-2003 (TREMBlrel. 23, Created)	
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)	
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)	
DE Hypothetical protein (fragment).	
GN PFA0655W.	
OS Plasmodium falciparum (isolate 3D7).	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX NCBI_TaxID=36329;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=22255708; PubMed=12368867;	
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,	
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,	
RA Buckee C.O., Burrows C., Chervach I., Chillingworth C.,	
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,	
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,	
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,	
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,	
RA Humphray S., Jagals K., James K.D., Johnson D., Kerhornou A.,	
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,	
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,	
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitz E.,	
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,	
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,	
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,	
RA Tustlon J.E., Craig A., Newbold C., Barrell B.G;	
RL "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13.";	
RT Nature 419:527-531(2002).	
DR EMBL; AL031746; CAD49077.1; -.	
KW Hypothetical protein.	
FT NON TER 851	
SQ SEQUENCE 851 AA; 105268 MW; 76ABED8CFCEB9001 CRC64;	
Query Match	30.7%; Score 59; DB 5; Length 851;
Best Local Similarity	29.0%; Pred.No.47;
Matches	9; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
OY	2 KKTFLSVHEVMVAENVAKDSLQWITDQYNKE 32 ::: ::: ::: ::: ::: ::: ::: ::: :::
Db	742 KQSMINVDDNFIRKYDDQKFQWIDENEKD 772 ::: ::: ::: ::: ::: ::: ::: ::: :::
RESULT 7	
Q81D83	PRELIMINARY; PRT; 2030 AA.
ID Q81D83	
AC Q81D83;	
DT 01-MAR-2003 (TREMBlrel. 23, Created)	
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)	
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)	
DE MALL3P1.310 protein.	
GN MALL3P1.310.	
OS Plasmodium falciparum (isolate 3D7).	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX NCBI_TaxID=36329;	
RN [1]	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 21.5714 Seconds
(without alignments)

526.560 Million cell updates/sec

Title: US-09-941-314-6

Perfect score: 189

Sequence: 1 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp rvirus:*
16: sp bacteriaph:*
17: sp archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	103	4 Q8WXU5	Q8WXU5 homo sapien
2	189	100.0	138	4 Q8WXU6	Q8WXU6 homo sapien
3	100	52.9	139	11 Q8K5A3	Q8K5A3 rattus norv
4	65	34.4	844	5 Q86S89	Q86S89 polyploastro
5	58	30.7	240	16 Q8XBD0	Q8XBD0 escherichia
6	58	30.7	851	5 Q81225	Q81225 plasmodium
7	58	30.7	2030	5 Q81D83	Q81D83 plasmodium
8	57.5	30.4	184	16 Q927X9	Q927X9 chlamydia p
9	56	29.6	209	5 Q62173	Q62173 caenorhabdi
10	55	29.1	175	16 Q81J78	Q81J78 bacillus an
11	55	29.1	327	2 Q31093	Q31093 rhizobium l
12	55	29.1	343	16 Q8UD47	Q8UD47 agrobacteri
13	55	29.1	413	16 Q7VKX1	Q7VKX1 haemophilus
14	55	29.1	665	10 Q9W4E8	Q9W4E8 cucumis sat
15	55	29.1	668	10 Q9FSY7	Q9FSY7 corylus ave
16	54.5	28.8	273	17 Q8Z245	Q8Z245 pyrobaculum

17	54.5	28.8	2226	5	Q97225	Q97225 plasmodium
18	54	28.6	176	16	Q814Q9	Q814Q9 bacillus ce
19	54	28.6	189	16	Q8A4V4	Q8A4V4 bacteroides
20	54	28.6	524	10	Q9SBN2	Q9SBN2 volvox cart
21	53.5	28.3	500	16	Q8AB15	Q8AB15 bacteroides
22	53	28.0	194	5	Q61819	Q61819 caenorhabdi
23	53	28.0	205	5	Q8T7Y7	Q8T7Y7 caenorhabdi
24	53	28.0	208	16	Q7UAB1	Q7UAB1 shigella fl
25	53	28.0	212	16	Q83R45	Q83R45 shigella fl
26	53	28.0	244	16	Q8FGM5	Q8FGM5 escherichia
27	53	28.0	300	5	Q86KT1	Q86KT1 dictyostell
28	53	28.0	351	2	Q48418	Q48418 klebsiella
29	53	28.0	351	2	Q8GFR0	Q8GFR0 citrobacter
30	53	28.0	543	9	Q859F0	Q859F0 pseudomonas
31	52	27.5	321	2	Q8G8T6	Q8G8T6 pseudomonas
32	52	27.5	404	10	Q9SY99	Q9SY99 arabidopsis
33	52	27.5	457	10	Q9SUX4	Q9SUX4 arabidopsis
34	52	27.5	592	3	Q8X228	Q8X228 mycosphaere
35	52	27.5	1227	5	Q9BIA2	Q9BIA2 caenorhabdi
36	52	27.5	1422	5	O00914	O00914 plasmodium
37	51.5	27.2	122	15	Q8EAA5	Q8EAA5 human immun
38	51.5	27.2	318	16	Q8X714	Q8X714 escherichia
39	51.5	27.2	465	16	Q9FJ81	Q9FJ81 campylobact
40	51.5	27.2	567	17	O26602	O26602 methanobact
41	51.5	27.2	656	5	Q24842	Q24842 entamoeba h
42	51	27.0	190	16	Q97KN6	Q97KN6 clostridium
43	51	27.0	216	3	Q05435	Q05435 saccharomyc
44	51	27.0	238	16	Q8CN34	Q8CN34 staphylococ
45	51	27.0	326	1	Q9UXP0	Q9UXP0 methanoblob

ALIGNMENTS

RESULT 1
Q8WXU5 ID Q8WXU5 PRELIMINARY; PRT; 103 AA.
AC Q8WXU5;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE SC13delta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335481; AAL71992.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 100.0%; Score 189; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.4e-17;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 36
|||||
Db 27 RKKTFLSVHEVAVENYAKDSLQWITDQYNKESDDK 62

RESULT 2
Q8WXU6 ID Q8WXU6 PRELIMINARY; PRT; 138 AA.
AC Q8WXU6;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)

OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92361242; PubMed=1822990;
 RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
 RT "The tobacco luminal binding protein is encoded by a multigene
 family";
 RL Plant Cell 3:1025-1035(1991).
 RN [2]
 RN ERRATUM.
 RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
 RL Plant Cell 3:1251-1251(1991).
 CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
 CC multimeric protein complexes inside the ER.
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
 CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60059; CAA42661.1; -;
 DR PIR; P02622; S21878.
 DR HSSP; P08109; 1CKR.
 DR InterPro; IPR000886; ER_target_S.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR PRINTS; PR00301; HEATSHOCK70.
 DR ProDom; PD000089; Hsp70; 1.
 DR PROSITE; PS00297; HSP70_1; PARTIAL.
 DR PROSITE; PS00329; HSP70_2; PARTIAL.
 DR PROSITE; PS01036; HSP70_3; PARTIAL.
 DR PROSITE; PS00014; ER_TARGET; 1.
 KW ATP-binding; Endoplasmic reticulum; Multigene family.
 FT NON_TER 1 1
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 289 292 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 292 AA; 32260 MW; 6E7A4F5107C6E2D5 CRC64;
 Query Match 27.0%; Score 51; DB 1; Length 292;
 Best Local Similarity 30.0%; Pred. No. 18;
 Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
 Db 219 LESDEKEKIEATATKEALEWLDNQSAEKD 248
 RESULT 15
 BIP8 TOBAC STANDARD; PRT; 293 AA.
 AC Q03686;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Luminal binding protein 8 (Bip 8) (78 kDa glucose-regulated protein
 DE hemolog 8) (GRP 78-8) (Fragment).
 GN BIP8.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RN SEQUENCE FROM N.A.

RX MEDLINE=92361242; PubMed=1822990;
 RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
 RT "The tobacco luminal binding protein is encoded by a multigene
 family";
 RL Plant Cell 3:1025-1035(1991).
 RN [2]
 RN ERRATUM.
 RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
 RL Plant Cell 3:1251-1251(1991).
 CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
 CC multimeric protein complexes inside the ER.
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
 CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60062; CAA42664.1; -;
 DR PIR; S21881; S21881.
 DR HSSP; P08109; 1CKR.
 DR InterPro; IPR000886; ER_target_S.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR PRINTS; PR00301; HEATSHOCK70.
 DR ProDom; PD000089; Hsp70; 1.
 DR PROSITE; PS00297; HSP70_1; PARTIAL.
 DR PROSITE; PS00329; HSP70_2; PARTIAL.
 DR PROSITE; PS01036; HSP70_3; PARTIAL.
 DR PROSITE; PS00014; ER_TARGET; 1.
 KW ATP-binding; Endoplasmic reticulum; Multigene family.
 FT NON_TER 1 1
 FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 290 293 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 293 AA; 32342 MW; 44939595E45650E6 CRC64;
 Query Match 27.0%; Score 51; DB 1; Length 293;
 Best Local Similarity 30.0%; Pred. No. 18;
 Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
 Db 222 LESDEKEKIEATATKEALEWLDNQSAEKD 251
 Search completed: March 18, 2004, 14:16:10
 Job time : 5.28571 secs

```

QY 1 RKTFLSVHEVMAVENYAKDSLOWITDQYNKESDD 35
DB 52 RTLYLSVEE---QDNLQEARLWLCYHWNLEDD 83

RESULT 12
BIP1_TOBAC STANDARD; PRT; 168 AA.
AC Q03683;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 3 (BIP 3) (78 kDa glucose-regulated protein
DE homolog 3) (GRP 78-3) (Fragment).
GN BIP3.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RP SEQUENCE FROM N.A.
RX MEDLINE=92361242; PubMed=1822990;
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RT "The tobacco luminal binding protein is encoded by a multigene
RT family.";
RL Plant Cell 3:1025-1035(1991).
RN [2]
RP ERRATUM.
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RL Plant Cell 3:1251-1251(1991).
CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60061; CAA42663.1; -
CC HSP; P04475; 1DKY.
CC InterPro; IPR000886; ER_target_S.
CC Pfam; PF00012; HSP70.
CC ProDom; PD000089; HSP70; 1.
CC PROSITE; PS00297; HSP70_1; PARTIAL.
CC PROSITE; PS00329; HSP70_2; PARTIAL.
CC PROSITE; PS01036; HSP70_3; PARTIAL.
CC PROSITE; PS00014; ER_TARGET; 1.
CC ATP-binding; Endoplasmic reticulum; Multigene family.
CC NON_TER 1
CC CARBOHYD 120 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SITE 165 168 PREVENT SECRETION FROM ER (POTENTIAL).
CC SEQUENCE 168 AA; 18838 MW; D0089CF2219C624E CRC64;

Query Match 27.0%; Score 51; DB 1; Length 168;
Best Local Similarity 30.0%; Pred. No. 9.6;
Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
DB 98 LESDEKEKETATKEALEWLDNDSAEKED 127

RESULT 13
BIP1_TOBAC STANDARD; PRT; 290 AA.
AC Q03682;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 2 (BIP 2) (78 kDa glucose-regulated protein
DE homolog 2) (GRP 78-2) (Fragment).
GN BIP2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RP SEQUENCE FROM N.A.
RX MEDLINE=92361242; PubMed=1822990;
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RT "The tobacco luminal binding protein is encoded by a multigene
RT family.";
RL Plant Cell 3:1025-1035(1991).
RN [2]
RP ERRATUM.
RA Denecke J., Goldman M.H., Demolder J., Seurinck J., Botterman J.;
RL Plant Cell 3:1251-1251(1991).
CC -!- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60060; CAA42662.1; -
CC PIR; S21877; S21877.
CC HSP; P08109; 1CKR.
CC InterPro; IPR000886; ER_target_S.
CC InterPro; IPR001023; HSP70.
CC Pfam; PF00012; HSP70; 1.
CC PRINTS; PR00301; HEATSHOCK70.
CC ProDom; PD000089; HSP70; 1.
CC PROSITE; PS00297; HSP70_1; PARTIAL.
CC PROSITE; PS00329; HSP70_2; PARTIAL.
CC PROSITE; PS01036; HSP70_3; PARTIAL.
CC PROSITE; PS00014; ER_TARGET; 1.
CC ATP-binding; Endoplasmic reticulum; Multigene family.
CC NON_TER 1
CC CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SITE 287 290 PREVENT SECRETION FROM ER (POTENTIAL).
CC SEQUENCE 290 AA; 32031 MW; E80F65136B64BF89 CRC64;

Query Match 27.0%; Score 51; DB 1; Length 290;
Best Local Similarity 30.0%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLOWITDQYNKESDD 35
DB 219 LESDEKEKETATKEALEWLDNDSAEKED 248

RESULT 14
BIP2_TOBAC STANDARD; PRT; 292 AA.
AC Q03682;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 2 (BIP 2) (78 kDa glucose-regulated protein
DE homolog 2) (GRP 78-2) (Fragment).
GN BIP2.

```



```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5860;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95059225; PubMed=7969277;
RA Cheng Q., Saul A.;
RT "The dihydrofolate reductase domain of rodent malarial: point mutations and pyrimethamine resistance."
RL Mol. Biochem. Parasitol. 65:361-363(1994).
CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADPH(+) = 7,8-dihydrofolate + NADPH.
CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP = dihydrofolate + dUMP.
CC -1- PATHWAY: Essential step for de novo glycine and purine synthesis, DNA precursor synthesis, and for the conversion of dUMP to dTMP.
CC -1- SIMILARITY: In the N-terminal section; belongs to the dihydrofolate reductase family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the thymidylate synthase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28121; AAA29582.1; -.
DR HSSP; P00379; 1RA9.
DR InterPro; IPR001796; DHFR.
DR ProSite; PS000398; Thymidylat_synth.
DR Pfam; PF00186; Dihfolate_red; 1.
DR PRINTS; PR00070; DHFR.
DR PROSITE; PS00075; DHFR; 1.
DR PROSITE; PS00091; THYMIDYLATE SYNTHASE; FALSE_NEG.
KW Multifunctional enzyme; Oxidoreductase; Transferase; NADPH; Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.
FT NON TER 1 182
FT DOMAIN <1> >182 DIHYDROFOLATE REDUCTASE.
FT NON TER 182 182
FT SEQUENCE 182 AA; 20924 MW; 33CD27441763198 CRC64;
SQ
Query Match 27.5%; Score 52; DB 1; Length 182;
Best Local Similarity 38.2%; Pred. No. 7.7;
Matches 13; Conservative 7; Mismatches 12; Indels 2; Gaps 1;
QY 3 KTFSLVHEWAVENYAKDSLQWITDQYNKESDDK 36
DB 42 KHFVSVTSYNNENYIR--LKWKRDKYIKENNVK 73
RESULT 11
HEM1_SALTY STANDARD; PRT; 418 AA.
ID AC P13581;
DT 01-JAN-1990 (Rel. 13, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.1-) (Glutr).
GN HEMA OR STM1777 OR STY1902 OR T1099.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhimurium; STRAIN=LT2; MEDLINE=89291746; PubMed=2544564;
RA Elliott T.;
RT "Cloning, Genetic characterization, and nucleotide sequence of the hemA-prfA operon of Salmonella typhimurium."
J. Bacteriol. 171:3948-3960(1989).
[2]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720; MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2."
RL Nature 413:852-856(2001).
[3]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhi; STRAIN=CT18; MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Baker S., Basham D., Brooks K., Chillingworth T., Connor P., Cronin A., Davies P., Davies R.M., Dowd L., White N., Farrar J., Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K., Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C., Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
[4]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931; MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J., Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18."
J. Bacteriol. 185:2330-2337(2003).
CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-semialdehyde + NADP(+) + tRNA(Glu).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -1- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04243; AAA88610.1; -.
DR EMBL; A8008779; AAL20692.1; -.
DR EMBL; AL627271; CAD02132.1; -.
DR EMBL; A8016837; AAC068762.1; -.
DR PIR; A32661; BVEBHA.
DR HSSP; Q42843; 1B29.
DR HAMAP; MG00087; hemA.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr_dimer; 1.
DR Pfam; PF05201; Glutr_N; 1.
DR Pfam; PF05200; Glutr_NAD_bind; 1.
DR TIGRFAMs; TIGR01035; hemA; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT ACT_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 99 99 BASE (BY SIMILARITY).
FT CONFLICT 139 139 E -> R (IN REF. 1).
SQ SEQUENCE 418 AA; 46105 MW; 92961964C36AD3F4 CRC64;
Query Match 27.2%; Score 51.5; DB 1; Length 418;
Best Local Similarity 31.4%; Pred. No. 23;
Matches 11; Conservative 8; Mismatches 13; Indels 3; Gaps 1;

```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=86176730; PubMed=3515318;
RA Sharma S., Stark T.F., Beattie W.G., Moses R.E.;
RT "Multiple control elements for the uvrC gene unit of Escherichia
RL coli.";
RN Nucleic Acids Res. 14:2301-2318(1986).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97251358; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampaio G., Seki Y., Sivasubraman S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 40.1-50.0 min region on the linkage map.";
RN DNA Res. 3:379-392(1996).
[4]
RP FUNCTION.
RX MEDLINE=92007787; PubMed=1915297;
RA Wang X., de Boer P.A.J., Rothfield L.I.;
RT "A factor that positively regulates cell division by activating
RT transcription of the major cluster of essential cell division genes
RL of Escherichia coli.";
RN EMBO J. 10:3363-3372(1991).
CC -I- FUNCTION: ACTIVATES CELL DIVISION BY SPECIFICALLY INCREASING
CC TRANSCRIPTION FROM ONE OF THE TWO PROMOTERS THAT LIE IMMEDIATELY
CC UPSTREAM OF THE FTSQAZ GENE CLUSTER.
CC -I- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X03691; CAA27327.1; -.
DR EMBL; AE000284; AAC74983.1; -.
DR EMBL; D90832; BAA15736.1; -.
DR PIR; A64955; Q0BCU1.
DR EcGene; EG10935; sdia.
DR InterPro; IPR005143; Autoind_bind.
DR InterPro; IPR000792; HTH_LuxR.
DR Pfam; PF03472; Autoind_bind; 1.
DR Pfam; PF00196; GerE; 1.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR PROSITE; PS00622; HTH_LUXR_FAMILY; 1.
KW Transcription regulation; Activator; DNA-binding; Trans-acting factor;
KW Cell division; Complete proteome.
FT DNA BIND 197 216 H-T-H MOTIF (BY SIMILARITY).
FT CONFLICT 120 128 TQVLMPLNR -> HSNVNAQTG (IN REF. 1).
FT CONFLICT 140 140 A -> R (IN REF. 1).
SQ SEQUENCE 240 AA; 28117 MW; C04CCC50C06135C4 CRC64;
Query Match 28.0%; Score 53; DB 1; Length 240;
Best Local Similarity 31.0%; Pred. No. 7.7;
Matches 9; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 3 KTFSLSHVEYMAVENYAKDSLOWITDQYNK 31
| | | | | | | | | | | | | | | | |
Db 74 KNFLAIDPVLNPFNSQGHLMWDDLFSE 102
| | | | | | | | | | | | | | | | |

RESULT 9
PREI STAAU STANDARD; PRT; 403 AA.
AC P03857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Plasmid recombination enzyme (Mobilization protein).
GN PRE OR MOB.
OS Staphylococcus aureus.
OG Plasmid pE194.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RX MEDLINE=82167187; PubMed=6279574;
RA Horinouchi S., Weisblum B.;
RT "Nucleotide sequence and functional map of pE194, a plasmid that
RT specifies inducible resistance to macrolide, lincosamide, and
RT streptogramin type B antibiotics.";
RL J. Bacteriol. 150:804-814(1982).
CC -I- FUNCTION: THE INTERACTION OF THE RSA SITE AND THE PRE PROTEIN MAY
CC NOT ONLY SERVES A FUNCTION IN PLASMID MAINTENANCE, BUT MAY ALSO
CC CONTRIBUTE TO THE DISTRIBUTION OF SMALL ANTIBIOTIC RESISTANCE
CC PLASMIDS AMONG GRAM-POSITIVE BACTERIA.
CC -I- MISCELLANEOUS: PRE PROTEINS CONTAIN CONSERVED POSITIVELY CHARGED
CC AMINO ACIDS PROBABLY INVOLVED IN THE BINDING OF THE PRE PROTEIN TO
CC THE RSA SITE.
CC -I- SIMILARITY: TO OTHER PRE PROTEINS (FROM PLASMIDS PUB110, PMV158,
CC PE194, PT181, PT8913), IN THEIR N-TERMINAL ONLY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; V01278; CAA24593.1; -.
DR InterPro; IPR001668; Mob_Pre.
DR Pfam; PF01076; Mob_Pre; 1.
DR Plasmid; DNA-binding.
FT BINDING 45 45 DNA (POTENTIAL).
FT BINDING 115 115 DNA (POTENTIAL).
SQ SEQUENCE 403 AA; 47839 MW; 875867A394000FC1 CRC64;
Query Match 28.0%; Score 53; DB 1; Length 403;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 8; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 14 VENVAKDSLOWITDQYNKES 33
| | | | | | | | | | | | | | |
Db 100 IERFPKDSLEFLENEYGKEN 119
| | | | | | | | | | | | | | | | |

RESULT 10
DRTS PLAVN STANDARD; PRT; 182 AA.
AC P46103;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS)
DE [includes: Dihydrofolate reductase (EC 1.5.1.3); Thymidylate synthase
DE (EC 2.1.1.45)] (Fragment).
OS Plasmodium vinckei.
```

```
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 663 LUMINAL BINDING PROTEIN 2.
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 660 663 PREVENT SECRETION FROM ER.
SQ SEQUENCE 663 AA; 73084 MW; 676987D19A0D918A CRC64;

Query Match 28.6%; Score 54; DB 1; Length 663;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLQWITDQYNKESDD 35
DB 592 LEAEKEKVEEALKEALEWLDNDSQAEKED 621

RESULT 6
BIP3_MAIZE
ID BIP3_MAIZE STANDARD; PRT; 663 AA.
AC O24581;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luminal binding protein 3 precursor (BiP3).
GN BIPE3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Flourey-2; TISSUE=Kernel;
RX MEDLINE=98094260; PubMed=9434171;
RA Wobrel R.L., Obrian G.R., Boston R.S.;
RT "Comparative analysis of BiP gene expression in maize endosperm.";
RL Gene 204:105-113(1997).
CC -1- FUNCTION: Probably plays a role in facilitating the assembly of
CC multimeric protein complexes inside the ER.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U58209; AAC49900.1; -.
DR PIR; T04080; T04080.
DR HSP; P19120; 3HSC.
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
KW ATP-binding; Endoplasmic reticulum; Multigene family; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 663 LUMINAL BINDING PROTEIN 3.
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 660 663 PREVENT SECRETION FROM ER.
SQ SEQUENCE 663 AA; 73156 MW; A710278D9F692723 CRC64;

Query Match 28.6%; Score 54; DB 1; Length 663;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLQWITDQYNKESDD 35
DB 592 LEAEKEKVEEALKEALEWLDNDSQAEKED 621
```

```
DB 592 LEAEKEKVEEALKEALEWLDNDSQAEKED 621

RESULT 7
MLR_LUMTE
ID MLR_LUMTE STANDARD; PRT; 195 AA.
AC P80164;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myosin regulatory light chain, striated muscle, 25 kDa isoform (LC25).
OS Lumbricus terrestris (Common earthworm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Lumbricus.
OX NCBI_TaxID=6398;
RN [1]
SEQUENCE.
RX MEDLINE=93145966; PubMed=8425543;
RA Serwe M., Meyer H.E., Craig A.G., Carlhoff D., D'Haese J.;
RT "Complete amino acid sequence of the regulatory light chain of
RT obliquely striated muscle myosin from earthworm, Lumbricus
RT terrestris.";
RL Eur. J. Biochem. 211:341-346(1993).
CC -1- FUNCTION: Plays an important role in regulation of muscle cell
CC contractile activity.
CC -1- SUBUNIT: Myosin is an hexamer of 2 heavy chains and 4 light
CC chains.
CC -1- MISCELLANEOUS: This chain binds calcium.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
DR PIR; S28845; S28845.
DR HSP; P13543; 1SCW.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 2.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; Eph; 2.
DR PROSITE; PS00018; EF_HAND; 1.
KW Myosin; Calcium-binding; Muscle protein.
FT DOMAIN 2 14 LYS-RICH (BASIC).
FT CA BIND 68 79 EF-HAND (BY SIMILARITY).
FT VARIANT 65 65 T -> A (IN 20% OF THE MOLECULES).
FT VARIANT 70 70 D -> N (IN 20% OF THE MOLECULES).
FT VARIANT 76 78 GPD -> DES (IN 20% OF THE MOLECULES).
FT VARIANT 81 82 GN -> AA (IN 20% OF THE MOLECULES).
FT VARIANT 84 84 F -> Y (IN 20% OF THE MOLECULES).
FT VARIANT 127 128 GT -> SS (IN 20% OF THE MOLECULES).
FT VARIANT 141 141 L -> V (IN 20% OF THE MOLECULES).
FT VARIANT 168 168 V -> L (IN 20% OF THE MOLECULES).
SQ SEQUENCE 195 AA; 21945 MW; 34BF85EB4D4E7F73 CRC64;

Query Match 28.0%; Score 53; DB 1; Length 195;
Best Local Similarity 52.6%; Pred. No. 6.2;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 14 VENYAKDSLQWITDQYNKE 32
DB 145 LEEYVXDLTLNVGDQYNKD 163

RESULT 8
SDIA_ECOLI
ID SDIA_ECOLI STANDARD; PRT; 240 AA.
AC P07026; P76313;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE regulatory protein sdia.
GN SDIA OR B1916.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Bacterichia.
OX NCBI_TaxID=562;
RN [1]
```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AL096677; CAC13170.1; -
 DR EMBL; AL096677; CAC17423.1; -
 DR HSSP; P01038; IAG0.
 DR Genew; HGNC:15959; CST11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KW Thiol protease inhibitor; Signal; Alternative splicing.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 137
 FT SITE 75 79 CYSSTATIN 11.
 FT SITE 93 101 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 114 134 BY SIMILARITY.
 FT DISULFID 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 76 110 Missing (in isoform 2).
 FT VARSPLIC 76 110 /FTID-VSP 001260.
 FT SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 189; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 3.6e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDDK 36
 DB 26 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDDK 61

RESULT 2

CS11_MOUSE STANDARD; PRT; 139 AA.
 ID CS11_MOUSE
 AC Q9D269;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin 11 precursor.
 GN CST11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 FT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AK020300; BAB32061.1; -
 DR HSSP; P01034; IG96.
 DR MGD; MGI:1925490; Cst11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KW Thiol protease inhibitor; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 139 CYSSTATIN 11.
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 94 102 BY SIMILARITY.
 FT DISULFID 115 135 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;

Query Match 55.0%; Score 104; DB 1; Length 139;
 Best Local Similarity 45.7%; Pred. No. 7.1e-07;
 Matches 16; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

QY 1 RKKTFLSVHVMVAVENYAKDSLQWITDQYNKESDD 35
 DB 27 KRKTFRIBEVSALESVKTLEYVTDEYNKESD 61

RESULT 3

MESJ_HAEIN STANDARD; PRT; 430 AA.
 ID MESJ_HAEIN
 AC P44689;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative cell cycle protein mesj homolog.
 GN MESJ OR H10404.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.P.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd."
 RL Science 269:496-512(1995).
 CC -!- SIMILARITY: BELONGS TO THE UPF0072 (MESJ/YCF62) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U32723; AAC22063.1; -

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 4.28571 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-6

Perfect score: 189

Sequence: 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	104	55.0	139	CS11 MOUSE	Q9d269 mus musculus
3	59.5	31.5	430	MESJ HABIN	P44689 haemophilus
4	55	29.1	257	FR12 SOYBN	Q94ic4 glycine max
5	54	28.6	663	BIP2 MAIZE	Z24067 zea mays (m
6	54	28.6	663	BIP3 MAIZE	Q24581 zea mays (m
7	53	28.0	195	MLR LUMTE	P80164 lumbricus t
8	53	28.0	240	SDIA ECOLI	P07026 escherichia
9	53	28.0	403	PRE1 STAAU	P03857 staphylococ
10	52	27.5	182	DRTS PLAVN	P46103 plasmodium
11	51.5	27.2	418	HEMA_SALTY	F13581 salmonella
12	51	27.0	168	BIP3 TOBAC	Q03683 nicotiana t
13	51	27.0	290	BIP1 TOBAC	Q03681 nicotiana t
14	51	27.0	292	BIP2 TOBAC	Q03682 nicotiana t
15	51	27.0	293	BIP8 TOBAC	Q03686 nicotiana t
16	51	27.0	326	MER_MERTI	Q9uxp0 methanolobu
17	51	27.0	438	GATD_PYRFU	Q8u0x0 pyrococcus
18	51	27.0	438	GATD_PYRHO	O59132 pyrococcus
19	51	27.0	666	BIP LYCES	P49118 lycopersico
20	51	27.0	667	BIP4 TOBAC	Q03684 nicotiana t
21	51	27.0	668	BIP5 TOBAC	Q03685 nicotiana t
22	51	27.0	668	BIP_SPTOL	Q42434 spinacia ol
23	51	27.0	838	OS94 MOUSE	F48722 mus musculus
24	51	27.0	1818	HMW2 MYCFN	P75471 mycoplasma
25	50	26.5	302	TX34 CAEEL	Q9nah2 caenorhabdi
26	49.5	26.2	418	HEM1 ECOL6	O8fi03 escherichia
27	49.5	26.2	439	RHGI HUMAN	Q07960 homo sapien
28	49.5	26.2	611	SNF1_CANGA	Q00372 candida gla
29	49.5	26.2	976	SCPI_HUMAN	Q15431 homo sapien
30	49	25.9	328	MER_METMA	O8p266 methanosarc
31	49	25.9	438	GATD_PYRAB	Q9v0t9 pyrococcus
32	49	25.9	474	SYE_VIBPA	O87rl6-vibrio para
33	49	25.9	484	SYE_STAAM	Q99w75 staphylococ

34	49	25.9	657	1	LONI THEAC	Q9hj89 thermoplas
35	49	25.9	918	1	IL6B HUMAN	P40189 homo sapien
36	48.5	25.7	326	1	FAH2_STAEP	O8cub8 staphylococ
37	48.5	25.7	504	1	GUNW_ERWCA	Q59395 erwinia car
38	48.5	25.7	792	1	PEPX_LACDL	P40334 lactobacill
39	48.5	25.7	793	1	PEPX_LACHE	Q59485 lactobacill
40	48.5	25.7	944	1	TPSX_SCHPO	O14081 schizosacch
41	48	25.4	324	1	Y089_BORBU	O51116 borrelia bu
42	48	25.4	326	1	XC00_PVRAB	Q9uz88 pyrococcus
43	48	25.4	348	1	VIPB_SALTY	Q04973 salmonella
44	48	25.4	426	1	CG2A_PATVU	P24861 patella vul
45	48	25.4	444	1	PST1_YEAST	Q12355 saccharomyc

ALIGNMENTS

RESULT 1
CS11_HUMAN STANDARD; PRT; 137 AA.
AC Q9H112; Q9H113;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Cystatin 11 precursor.
GN Cst11 OR CST8L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Corby N.R.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Dunn M.,
RA Coulson A., Coville G.J., Deadman R., Dhali P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehaeslaih M.H., Leveraha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whithead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;

"The DNA sequence and comparative analysis of human chromosome 20.";
Nature 414:865-871(2001).
RL - SUBCELLULAR LOCATION: Secreted (Potential).
CC - ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q9H112-1; Sequence=Displayed;
Name=2;
IsoId=Q9H112-2; Sequence=VSP_001260;
Notes=No experimental confirmation available;
Note=No experimental confirmation available;
CC - SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)

XX SQ Sequence 117 AA;
 Query Match 100.0%; Score 189; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.9e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
 |||||
 DB 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 41
 |||||

RESULT 15
 AAU79852
 ID AAU79852 standard; protein; 137 AA.
 XX
 AC AAU79852;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8).
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI: 2002-383044/41.
 DR N-PSDB; ABK49522.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 93-94; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)

XX SQ Sequence 137 AA;
 Query Match 100.0%; Score 189; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 3.5e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
 |||||
 DB 26 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 61
 |||||

Search completed: March 18, 2004, 14:14:58
 Job time : 33.6429 secs

CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemorrhoidosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;

Query Match 100.0%; Score 189; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36

RESULT 13
AAU79853
ID AAU79853 standard; protein; 115 AA.
XX
AC AAU79853;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8) antigenic fragment #1.
XX
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic fragment.
XX
OS Homo sapiens.
XX
FN WO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026868. *see in situ*
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the

CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic fragment of human cystatin-8
CC (Zcys8)
XX
SQ Sequence 115 AA;

Query Match 100.0%; Score 189; DB 5; Length 115;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36
DB 4 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 39

RESULT 14
AAU79854
ID AAU79854 standard; protein; 117 AA.
XX
AC AAU79854;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8) antigenic fragment #2.
XX
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic fragment.
XX
OS Homo sapiens.
XX
FN WO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 94-95; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKKTFLSVHVMVAENYAKDSLQWITDQYNKESDDK 36
|||||

Db 1 RKKTFLSVHVMVAENYAKDSLQWITDQYNKESDDK 36
|||||

RESULT 11

AA02833

ID AA02833 standard; protein; 50 AA.

XX

AC AA02833;

XX

DT 09-OCT-2001 (first entry)

XX

DE Peptide #1515 encoded by probe for measuring breast gene expression.

XX

KW Probe; human; breast disease; breast cancer; development disorder;

KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX

OS Homo sapiens.

XX

PN WO200157270-A2.

XX

PD 09-AUG-2001.

XX

PF 29-JAN-2001; 2001WO-US000661.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR

PR 26-MAY-2000; 2000US-0207456P.

PR

PR 30-JUN-2000; 2000US-00608408.

PR

PR 03-AUG-2000; 2000US-00632366.

PR

PR 21-SEP-2000; 2000US-0234687P.

PR

PR 27-SEP-2000; 2000US-0236359P.

PR

PR 04-OCT-2000; 2000GB-00024263.

XX

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

XX WPI; 2001-476286/51.

XX

XX Novel single exon nucleic acid probe used to measuring gene expression in

PT a human breast.

XX

XX Claim 27; SEQ ID NO 11573; 322pp; English.

XX

XX The present invention relates to novel single exon nucleic acid probes
CC (see AA00010-AA110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 50 AA;

Query Match 100.0%; Score 189; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 RKKTFLSVHVMVAENYAKDSLQWITDQYNKESDDK 36
|||||

Db

1 RKKTFLSVHVMVAENYAKDSLQWITDQYNKESDDK 36
|||||

RESULT 12

```
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
PT
XX
XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 50 AA;

    Query Match      100.0%; Score 189; DB 4; Length 50;
    Best Local Similarity 100.0%; Pred. No. 1e-18;
    Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
   |||||
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
   |||||

RESULT 9
AAM54871
ID AAM54871 standard; protein; 50 AA.
XX
AC AAM54871;
XX
XX 05-NOV-2001 (first entry)
DT
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 26976.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488946/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
PT
XX
XX Example 4; SEQ ID NO 26976; 650pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC
```

```
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 50 AA;

    Query Match      100.0%; Score 189; DB 4; Length 50;
    Best Local Similarity 100.0%; Pred. No. 1e-18;
    Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
   |||||
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
   |||||

RESULT 10
ABG48915
ID ABG48915 standard; peptide; 50 AA.
XX
XX ABG48915;
XX
XX 25-FEB-2003 (first entry)
DT
XX Human liver peptide, SEQ ID NO 27563.
DE
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488988/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
PT
XX
XX Claim 27; SEQ ID NO 27563; 658pp; English.
PS
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;

    Query Match      100.0%; Score 189; DB 4; Length 50;
    Best Local Similarity 100.0%; Pred. No. 1e-18;
```



```

XX AC AAM27545;
XX DT 17-OCT-2001 (first entry)
XX DE Peptide #1592 encoded by probe for measuring placental gene expression.
XX DE DE
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT Gene expression in human placenta.
XX PS Claim 27; SEQ ID NO 27814; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AAL31315-AAL57546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders
XX SQ Sequence 50 AA;
Query Match 100.0%; Score 189; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
RESULT 5
ABB32389
ID ABB32389 standard; peptide; 50 AA.
XX AC ABB32389;
XX DT 01-FEB-2002 (first entry)
XX DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.
XX DE DE
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX KW cancer.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.

```

```

XX PF 30-JAN-2001; 2001WO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a peptide
XX CC encoded by a single exon nucleic acid probe of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 50 AA;
Query Match 100.0%; Score 189; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
RESULT 6
ABB28913
ID ABB28913 standard; peptide; 50 AA.
XX AC ABB28913;
XX DT 01-FEB-2002 (first entry)
XX DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.
XX DE DE
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX KW cancer.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000662.

```

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 36 AA;

Query Match 100.0%; Score 189; DB 5; Length 36;
 Best Local Similarity 100.0%; Pred. No. 7.1e-19;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 2
 AAM15096
 ID AAM15096 standard; protein; 50 AA.
 AC AAM15096;
 XX
 DT 12-OCT-2001 (first entry)

DE Peptide #1530 encoded by probe for measuring cervical gene expression.
 XX Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO2001527278-A2.
 XX

09-AUG-2001,
 30-JAN-2001; 2001WO-US0000670.
 04-FEB-2000; 2000US-01803112P.
 26-MAY-2000; 2000US-0207456P.
 30-JUN-2000; 2000US-00608408.
 03-AUG-2000; 2000US-00632366.
 21-SEP-2000; 2000US-0234687P.
 27-SEP-2000; 2000US-0236359P.
 04-OCT-2000; 2000GB-00024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX
 PS Claim 27; SEQ ID NO 19922; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes
 CC (SENPs; see AAM10068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 50 AA;
 Query Match 100.0%; Score 189; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 3
 ABB34086
 ID ABB34086 standard; peptide; 50 AA.
 XX
 AC ABB34086;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #1592 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 OS
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US0000669.
 XX
 PR 04-FEB-2000; 2000US-01803112P.
 26-MAY-2000; 2000US-0207456P.
 30-JUN-2000; 2000US-00608408.
 03-AUG-2000; 2000US-00632366.
 21-SEP-2000; 2000US-0234687P.
 27-SEP-2000; 2000US-0236359P.
 04-OCT-2000; 2000GB-00024263.
 (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human foetal liver.
 XX
 PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human foetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 50 AA;
 Query Match 100.0%; Score 189; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 4
 AAM27545
 ID AAM27545 standard; protein; 50 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 32.6429 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-6
Perfect score: 189
Sequence: 1 RKKTFSLVHEVAVENYAKSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	189	100.0	36	5 AAU79856	Human cys
2	189	100.0	50	4 AAU15096	Peptide #
3	189	100.0	50	4 ABB34086	Peptide #
4	189	100.0	50	4 AAM27545	Peptide #
5	189	100.0	50	4 ABB32389	Peptide #
6	189	100.0	50	4 ABB28913	Peptide #
7	189	100.0	50	4 ABB19524	Protein #
8	189	100.0	50	4 AAM67252	Human bon
9	189	100.0	50	4 AAM54871	Human bra
10	189	100.0	50	4 ABG48915	Human liv
11	189	100.0	50	4 AAM02833	Peptide #
12	189	100.0	50	5 ABG36903	Human pep
13	189	100.0	115	5 AAU79853	Human cys
14	189	100.0	117	5 AAU79854	Human cys
15	189	100.0	137	5 AAU79852	Human cys
16	101	53.4	22	5 AAU79855	Human cys
17	99	52.4	18	5 AAU79857	Human cys
18	99	52.4	35	5 AAU79858	Human cys
19	59.5	31.5	430	6 ABB30233	Protein e
20	57.5	30.4	184	5 ABB30713	Chlamydia
21	57	30.2	623	7 ADC95099	S. faeciu
22	56.5	29.9	97	6 ABM73469	Staphyloc
23	56.5	29.9	579	6 ABM16559	Protein e
24	55	29.1	27	5 AAU79859	Human cys
25	55	29.1	49	5 AAU79863	Human cys

ALIGNMENTS

RESULT 1
AAU79856
ID AAU79856 standard; peptide; 36 AA.
XX
AC AAU79856;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8) antigenic fragment #4.
XX
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic fragment.
XX
OS Homo sapiens.
XX
PN WO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 95; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large

AbB09526 Soybean f
Abu44069 Protein e
Aay35198 Chlamydia
Abu28659 Protein e
Aar10545 Recombina
Abu44692 Protein e
Adc01083 Enterohae
Abu47755 Protein e
Aag82725 S. epider
Abp39544 Staphyloc
Aay00873 Original
Aay00870 S. tubero
AbB57153 Mouse isc
Abu36321 Protein e
Abu50049 Protein e
Aag22389 Arabidops
Aag51291 Arabidops
Ada36895 Acinetoba
Aag51290 Arabidops
Aag22388 Arabidops

26 55 29.1 256 5 ABB09526
27 54 28.6 239 6 ABU44069
28 53.5 28.3 195 2 AAY35198
29 53 28.0 240 6 ABU28659
30 52 27.5 918 2 AAR10545
31 51.5 27.2 278 6 ABU44692
32 51.5 27.2 318 7 ADC01083
33 51.5 27.2 418 6 ABU47755
34 51 27.0 192 4 AAG82725
35 51 27.0 277 5 ABP39544
36 51 27.0 458 2 AAY00873
37 51 27.0 766 2 AAY00870
38 51 27.0 838 5 ABB57153
39 51 27.0 1818 6 ABU36321
40 50.5 26.7 460 6 ABU50049
41 50 26.5 391 3 AAG22389
42 50 26.5 391 3 AAG51291
43 50 26.5 404 6 ADA36895
44 50 26.5 409 3 AAG51290
45 50 26.5 409 3 AAG22388

THIS PAGE BLANK (USPTO)

```

US-10-424-599-249999

Query Match      29.4%; Score 55.5; DB 12; Length 131;
Best Local Similarity 37.9%; Pred.No. 9;
Matches 11; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

Qy   3 KTFSLSHVEHVAENVY-AKOSLQWITDQYN 30
    :|::||::||::||::||::||::||::||:
Db   54 RLFADHSGEAAQTYNADDTLQWIWNQH 82


RESULT 14
US-09-941-314-9
; Sequence 9, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mamalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-9

Query Match      29.1%; Score 55; DB 9; Length 27;
Best Local Similarity 100.0%; Pred.No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   27 DQYNKESDDK 36
    |||||||
Db   1 DQYNKESDDK 10


RESULT 15
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mamalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-13

Query Match      29.1%; Score 55; DB 9; Length 49;
Best Local Similarity 100.0%; Pred.No. 3.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   27 DQYNKESDDK 36
    |||||||
Db   1 DQYNKESDDK 10


Search completed: March 18, 2004, 14:30:53
Job time : 22.7857 secs
```



```

Db      1 KOSLOWITDQYNKESDDK 18

RESULT 9
US-09-941-314-8
; Sequence 8, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-8

Query Match      52.4%; Score 99; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 KOSLOWITDQYNKESDDK 36
Db      1 KOSLOWITDQYNKESDDK 18

RESULT 10
US-10-282-122A-58157
; Sequence 58157, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308

Db      1 KOSLOWITDQYNKESDDK 18

Query Match      52.4%; Score 99; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 KOSLOWITDQYNKESDDK 36
Db      1 KOSLOWITDQYNKESDDK 18

RESULT 11
US-10-312-273-375
; Sequence 375, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035W0
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 375
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-375

Query Match      30.4%; Score 57.5; DB 15; Length 184;
Best Local Similarity 38.5%; Pred. No. 7;
Matches 15; Conservative 6; Mismatches 11; Indels 7; Gaps 1;

QY      1 RKKTFLSVHEWMAVENYAKD-----SLQWITDQYNKE 32
Db      26 RKKTFLSHRVLARPSFPVDYCPGKIYDLQEIYEELNAQ 64

RESULT 12
US-10-282-122A-44483
; Sequence 44483, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith

```

```

US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-3

Query Match      100.0%; Score 189; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 6.5e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
      |||||||
Db      4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 39
      |||||||

RESULT 5
US-09-941-314-4
; Sequence 4, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-4

Query Match      100.0%; Score 189; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.7e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
      |||||||
Db      6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 41
      |||||||

RESULT 6
US-09-941-314-2
; Sequence 2, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-2

Query Match      100.0%; Score 189; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 7.9e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
      |||||||
Db      26 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 61
      |||||||

RESULT 7
US-09-941-314-5
; Sequence 5, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-5

Query Match      53.4%; Score 101; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RKKTFLSVHEVMAVENYAKD 20
      |||||||
Db      3 RKKTFLSVHEVMAVENYAKD 22
      |||||||

RESULT 8
US-09-941-314-7
; Sequence 7, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-7

Query Match      52.4%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 KDSLQWITDQYNKESDDK 36
      |||||||

```

APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aomicca-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 34822
 LENGTH: 50
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AL109954.10
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.2
 OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
 OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
 US-09-864-761-48936

Query Match 100.0%; Score 189; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.6e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36
 DB 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36

RESULT 4

RESULT 3
 US-09-864-761-48936
 Sequence 48936, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aomicca-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 48936
 LENGTH: 50
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AL096677.18
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
 OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
 OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
 US-09-864-761-48936

Query Match 100.0%; Score 189; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.6e-18;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36
 DB 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDK 36

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 21.7857 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-6

Perfect score: 189

Sequence: 1 RKKTFSLVHEVAVENYAKDSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	36	9 US-09-941-314-6	Sequence 6, Appli
2	189	100.0	50	9 US-09-864-761-34822	Sequence 34822, A
3	189	100.0	50	9 US-09-864-761-48936	Sequence 48936, A
4	189	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
5	189	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
6	189	100.0	137	9 US-09-941-314-2	Sequence 2, Appli
7	101	53.4	22	9 US-09-941-314-5	Sequence 5, Appli
8	99	52.4	18	9 US-09-941-314-7	Sequence 7, Appli
9	99	52.4	35	9 US-09-941-314-8	Sequence 8, Appli
10	59.5	31.5	430	12 US-10-282-122A-58157	Sequence 58157, A
11	57.5	30.4	184	15 US-10-312-273-375	Sequence 375, App
12	56.5	29.9	579	12 US-10-282-122A-44483	Sequence 44483, A
13	55.5	29.4	131	12 US-10-424-599-249999	Sequence 249999,
14	55	29.1	27	9 US-09-941-314-9	Sequence 9, Appli
15	55	29.1	49	9 US-09-941-314-13	Sequence 13, Appli

Sequence 261063,	252	12	US-10-424-599-261063
Sequence 180261,	260	12	US-10-424-599-180261
Sequence 275467,	186	12	US-10-424-599-275467
Sequence 56041, A	369	12	US-10-425-114-56041
Sequence 275468,	486	12	US-10-424-599-275468
Sequence 57763, A	507	12	US-10-425-114-57763
Sequence 71993, A	239	12	US-10-282-122A-71993
Sequence 46564, A	684	12	US-10-425-114-46564
Sequence 616, App	195	15	US-10-289-762-616
Sequence 275473,	489	12	US-10-424-599-275473
Sequence 45919, A	494	12	US-10-425-114-45919
Sequence 56583, A	240	12	US-10-282-122A-56583
Sequence 72616, A	278	12	US-10-282-122A-72616
Sequence 64859, A	282	12	US-10-425-114-64859
Sequence 60692, A	384	12	US-10-425-114-60692
Sequence 70692, A	418	12	US-10-282-122A-70692
Sequence 72398, A	81	12	US-10-425-114-72398
Sequence 259133,	159	12	US-10-424-599-259133
Sequence 178326,	669	12	US-10-424-599-178326
Sequence 64245, A	1818	12	US-10-282-122A-64245
Sequence 77973, A	460	12	US-10-282-122A-77973
Sequence 62, Appli	430	15	US-10-320-800-62
Sequence 50269, A	469	12	US-10-282-122A-50269
Sequence 45567, A	512	12	US-10-425-114-45567
Sequence 58446, A	626	12	US-10-425-114-58446
Sequence 178328,	669	12	US-10-424-599-178328
Sequence 178322,	781	12	US-10-424-599-178322
Sequence 2949, Ap	1166	15	US-10-104-047-2949
Sequence 57562, A	1175	12	US-10-282-122A-57562
Sequence 10627, A	1208	9	US-09-815-242-10627

ALIGNMENTS

RESULT 1

US-09-941-314-6
; Sequence 6, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-6

Query Match 100.0%; Score 189; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RKKTFSLVHEVAVENYAKDSLQWITDQYNKESDDK 36
Db 1 RKKTFSLVHEVAVENYAKDSLQWITDQYNKESDDK 36

RESULT 2

US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.


```

Best Local Similarity 33.3%; Pred. No. 32;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLQWITDQYNKESDD 35
DB 592 LEAEKEKVEEALKEALEWLDNDQSAEKD 621
      | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | |

RESULT 10
T04080
dnak-type molecular chaperone cBiPe3 - maize
N;Alternate names: luminal binding protein cBiPe3
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: T04080
R;Wrobel, R.L.; O'Brian, G.R.; Boston, R.S.
Gene 204, 105-113, 1997
A;Title: Comparative analysis of BiP gene expression in maize endosperm.
A;Reference number: Z15191; MUID:98094260; PMID:9434171
A;Accession: T04080
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-663 <WRO>
A;Cross-references: EMBL:U58209; NID:g1575129; PIDN:AAC4900.1; PID:g1575130
C;Function:
A;Description: involved in protein folding and assembling/diseassembling of prot
C;Keywords: ATP; molecular chaperone

Query Match 28.6%; Score 54; DB 2; Length 663;
Best Local Similarity 33.3%; Pred. No. 32;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 LSVHEVMAVENYAKDSLQWITDQYNKESDD 35
DB 592 LEAEKEKVEEALKEALEWLDNDQSAEKD 621
      | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | |

RESULT 11
S28845
myosin regulatory light chain LC25, muscle - earthworm (Lumbricus terrestris)
C;Species: Lumbricus terrestris (common earthworm)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 18-Jun-1999
C;Accession: S28845
R;Serwe, M.; Meyer, H.E.; Craig, A.G.; Carlhoff, D.; d'Haese, J.
Eur. J. Biochem. 211, 341-346, 1993
A;Title: Complete amino acid sequence of the regulatory light chain of obliquel
A;Reference number: S28845; MUID:93145966; PMID:8425543
A;Accession: S28845
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-195 <SER>
C;Superfamily: calmodulin; calmodulin repeat homology
C;Keywords: EF hand
P;55-87/Domain: calmodulin repeat homology <EFL>

Query Match 28.0%; Score 53; DB 2; Length 195;
Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 14 VENVAKDSLQWITDQYNKE 32
DB 145 LEEYVKDLLTNVGDQYNKD 163
      : | | | | | | | | | |
      : | | | | | | | | | |

RESULT 12
T33117
hypothetical protein B0511.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jun-2000
C;Accession: T33117
R;Tin-Wollam, A.; Sutterer, C.; Ozersky, P.
submitted to the EMBL Data Library, May 1998

```

Best Local Similarity 34.5%; Pred. No. 2.9;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 KTFLSVHVEHVAENVYAKDSLQWITDQYNK 31
DB 74 KNFLAIDPVLNPFNSQGHLMWDDLFE 102

RESULT 3

G85808
hypothetical protein sdia [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85808
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001.
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Cross-references: GB:AE005174; NID:g12516014; PIDN:AAG56931.1; GSPDB:GN00145; UWGP:Z30
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: sdia
C:Superfamily: sdia regulatory protein

Query Match 30.7%; Score 58; DB 2; Length 240;
Best Local Similarity 34.5%; Pred. No. 2.9;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 KTFLSVHVEHVAENVYAKDSLQWITDQYNK 31
DB 74 KNFLAIDPVLNPFNSQGHLMWDDLFE 102

RESULT 4

D72061
hypothetical protein CP0175 [imported] - Chlamydia pneumoniae (strains CWL029 and AF
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: D72061; A81606
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72061
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <ARN>
A:Cross-references: GB:AE001642; GB:AE001363; NID:g4376865; PIDN:AAD18714.1; PID:g4376866
A:Experimental source: strain CWL029
R:Read, T.D.; Brunnham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81606
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <RNA>
A:Cross-references: GB:AE002178; GB:AE002161; NID:g7189099; PIDN:AAF38049.1; PID:g718910
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CPn0574; CP0175

Query Match 30.4%; Score 57.5; DB 2; Length 184;
Best Local Similarity 38.5%; Pred. No. 2.5;
Matches 15; Conservative 6; Mismatches 11; Indels 7; Gaps 1;

QY 1 RKKTFLSVHVEHVAENVYAKD-----SLOWITDQYNK 32

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 6.92857 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-6
Perfect score: 189
Sequence: 1 RKKTFLSVHEVMAVENYAKSLQWITDQYNKESDDK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Piri:*
2: Piri2:*
3: Piri3:*
4: Piri4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.5	31.5	430	2 D64151	hypothetical prote
2	58	30.7	240	2 F90960	hypothetical prote
3	58	30.7	240	2 G85808	hypothetical prote
4	57.5	30.4	184	2 D72061	hypothetical prote
5	57.5	30.4	184	2 B86562	hypothetical prote
6	56	29.6	209	2 T20975	hypothetical prote
7	55	29.1	343	2 AG2856	hypothetical prote
8	54	28.6	467	2 JQ0966	dnak-type molecula
9	54	28.6	663	2 T04078	dnak-type molecula
10	54	28.6	663	2 T04080	dnak-type molecula
11	53	28.0	195	2 S28845	myosin regulatory
12	53	28.0	205	2 T33117	hypothetical prote
13	53	28.0	240	1 Q0ECU1	probable transcrip
14	53	28.0	403	1 Q9SA4E	hypothetical prote
15	52	27.5	404	2 C96640	hypothetical prote
16	52	27.5	457	2 T05439	hypothetical prote
17	52	27.5	1422	2 T18404	chromatin remodel
18	51.5	27.2	318	2 C85517	unknown protein en
19	51.5	27.2	318	2 F90666	hypothetical prote
20	51.5	27.2	418	1 BV6BHA	glutamy1-trna redu
21	51.5	27.2	418	2 AC0720	glutamy1-trna redu
22	51.5	27.2	465	2 A81419	hypothetical prote
23	51.5	27.2	567	2 B69166	hypothetical prote
24	51.5	27.2	656	2 A48439	dnak-type molecula
25	51	27.0	190	2 F97008	hypothetical prote
26	51	27.0	290	2 S21877	dnak-type molecula
27	51	27.0	292	2 S21878	dnak-type molecula
28	51	27.0	293	2 S21881	dnak-type molecula
29	51	27.0	326	2 T45226	probable N5,N10-me

ALIGNMENTS

RESULT 1

D64151
hypothetical protein HI0404 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Dec-2002
C;Accession: D64151
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Godayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Guehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: D64151
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-430 <TIGR>
A;Cross-references: GB:U32723; GB:L42023; NID:gl573363; PIDN:AAC22063.1; PID:gl573375; TIG
A;Note: best homolog was a hypothetical protein from Bacillus subtilis
C;Superfamily: Cell cycle protein Mesaj

Query Match 31.5%; Score 59.5; DB 2; Length 430;
Best Local Similarity 52.4%; Pred. No. 3.6;
Matches 11; Conservative 6; Mismatches 3; Indels 1; Gaps 1;

Qy 14 VENYA-KDSLQWITDQYNKES 33

Db 169 LENYAQKEKLWITDSSNEDN 189

RESULT 2

F90960
hypothetical protein ECe2654 [imported] - Escherichia coli (strain O157:H7, substrain RIN
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90960
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90960
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA036077.1; PID:gl13362122; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
C;Superfamily: sdiA regulatory protein

Query Match 30.7%; Score 58; DB 2; Length 240;

Search completed: March 18, 2004, 14:25:35
Job time : 9.21429 secs

THIS PAGE BLANK (cont.)

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6509155 g312212
US-09-507-765-33

Query Match 26.2%; Score 49.5; DB 4; Length 439;
Best Local Similarity 39.4%; Pred. No. 41;
Matches 13; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 4 TELSVEHMAVENYAKDSLQWITDQYNKESDDK 36
DB 121 TLLYHLHGLTSDN--RPSLSWLRDAY-REFDRK 150

RESULT 10
US-09-198-452A-638
; Sequence 638, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:

; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 638
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae

US-09-198-452A-638

Query Match 26.2%; Score 49.5; DB 4; Length 460;
Best Local Similarity 35.9%; Pred. No. 44;
Matches 14; Conservative 10; Mismatches 8; Indels 7; Gaps 3;

QY 2 KKTFLSVHEVMAVENYAK--DSLQWITDQ---YNKESDD 35
DB 172 KXSLLEIRQVL--EDYAKLSAISFTQDKLWIEKESD 208

RESULT 11
US-09-392-714-24
; Sequence 24, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses Thereof
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-24

Query Match 26.2%; Score 49.5; DB 4; Length 973;
Best Local Similarity 37.8%; Pred. No. 1.1e+02;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 6 LSVHEV-----MAVENYAKDSLQWITDQYNKESDDK 36

DB 633 LNVYEIKVNKLELES-AKQKFGETTDYQKEIEDK 668

RESULT 12
US-09-104-324B-4
; Sequence 4, Application US/09104324B
; Patent No. 6232460
; GENERAL INFORMATION:
; APPLICANT: T recl, Ozlem; Sahin, Ugur; Pfreundschuh, Michael
; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers, And Methods For Identifying Pathogenic Markers In A Sample Of Tissue
; TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers In A Sample Of Tissue
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,324B
; FILING DATE: 25-June-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/892,702
; FILING DATE: 15-July-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6232460man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5491
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-104-324B-4

Query Match 26.2%; Score 49.5; DB 3; Length 976;
Best Local Similarity 37.8%; Pred. No. 1.1e+02;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 6 LSVHEV-----MAVENYAKDSLQWITDQYNKESDDK 36
DB 636 LNVYEIKVNKLELES-AKQKFGETTDYQKEIEDK 671

RESULT 13
US-08-913-578-2
; Sequence 2, Application US/08913578
; Patent No. 6218159
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6218159el tRNA synthetase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

Query Match	26.5%	Score 50;	DB 4;	Length 1224;
Best Local Similarity	35.3%	Pred. No. 1.2e+02;		

```

1  RESULT 9
2  US-09-507-765-33
3  ; Sequence 33, Application US/09507765
4  ; Patent No. 6509155
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Klingler, Elizabeth A.
7  ; APPLICANT: Stewart, Elizabeth A.
8  ; APPLICANT: Yue, Henry
9  ; APPLICANT: Baughn, Mariah R.
10 ; TITLE OF INVENTION: GTPASE ACTIVATING PROTEINS
11 ; FILE REFERENCE: PC-0010 US
12 ; CURRENT APPLICATION NUMBER: US/09/507,765
13 ; CURRENT FILING DATE: 2000-02-18
14 ; NUMBER OF SEQ ID NOS: 33
15 ; SOFTWARE: PERL Program
16 ; SEQ ID NO 33
17 ; LENGTH: 439
18 ; TYPE: PRT

```

```
Query Match      30.2%; Score 57; DB 4; Length 623;
Best Local Similarity 35.5%; Pred. No. 5.1;
Matches 11; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY      6 LSVHEVMAVENYAKDSLQWITDQYNKESDDK 36
DB      225 LKVEPMDEPGYSGEKIKETERYQKLADSK 255

RESULT 2
US-09-198-452A-616
; Sequence 616, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 616
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-616

Query Match      28.3%; Score 53.5; DB 4; Length 195;
Best Local Similarity 35.9%; Pred. No. 3.9;
Matches 14; Conservative 7; Mismatches 11; Indels 7; Gaps 1;

QY      1 RKKTFLSVHEVMAVENYAKD-----SLQWITDQYNKE 32
DB      37 RQKTFLSHRVLRAPSPDYVCPGKIYDLQEIYEELNAQ 75

RESULT 3
US-09-134-001C-4389
; Sequence 4389, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4389
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4389

Query Match      27.0%; Score 51; DB 4; Length 277;
Best Local Similarity 38.7%; Pred. No. 14;
Matches 12; Conservative 6; Mismatches 7; Indels 6; Gaps 1;

QY      1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNK 31
DB      204 RKKPFLSMPVPVAMYSL-----WLTKKKYKK 228

RESULT 4
US-09-463-238-15
; Sequence 15, Application US/09463238
; Patent No. 6469230
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
```

```
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Alison M
; APPLICANT: Bustos Guillen, Regla
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463,238
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/GB98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-463-238-15
```

```
Query Match      27.0%; Score 51; DB 4; Length 397;
Best Local Similarity 40.7%; Pred. No. 22;
Matches 11; Conservative 6; Mismatches 8; Indels 2; Gaps 2;

QY      9 HEVMAVENY-AKDSLQWITDQ-YNKES 33
DB      292 HNVLRKENFIEKNDITWLEDNWNYES 318
```

```
RESULT 5
US-09-463-238-6
; Sequence 6, Application US/09463238
; Patent No. 6469230
; GENERAL INFORMATION:
; APPLICANT: Edwards, Elizabeth A
; APPLICANT: Smith, Alison M
; APPLICANT: Bustos Guillen, Regla
; APPLICANT: Martin, Catherine R
; APPLICANT: Plant Bioscience Limited
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 97.118
; CURRENT APPLICATION NUMBER: US/09/463,238
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/GB98/02280
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: GB 9716185.5
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-463-238-6
```

```
Query Match      27.0%; Score 51; DB 4; Length 766;
Best Local Similarity 40.7%; Pred. No. 50;
Matches 11; Conservative 6; Mismatches 8; Indels 2; Gaps 2;

QY      9 HEVMAVENY-AKDSLQWITDQ-YNKES 33
DB      661 HNVLRKENFIEKNDITWLEDNWNYES 687
```

```
RESULT 6
US-09-328-352-8182
; Sequence 8182, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 9.21429 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-6
Perfect score: 189
Sequence: 1 RKKTFSLVHEWAVENYAKDSLOWITDQYNKESDDK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	30.2	623	4	US-09-107-532A-4726
2	53.5	28.3	195	4	US-09-198-452A-616
3	51	27.0	277	4	US-09-134-001C-4389
4	51	27.0	397	4	US-09-463-238-15
5	51	27.0	786	4	US-09-463-238-6
6	50	26.5	404	4	US-09-328-352-8182
7	50	26.5	1224	4	US-09-107-532A-6220
8	49.5	26.2	125	2	US-08-408-095-35
9	49.5	26.2	439	4	US-09-507-765-33
10	49.5	26.2	460	4	US-09-198-452A-638
11	49.5	26.2	973	4	US-09-392-714-24
12	49.5	26.2	976	3	US-09-104-324B-4
13	49	25.9	484	3	US-08-913-578-2
14	49	25.9	484	3	US-08-785-427-2
15	49	25.9	488	2	US-08-599-455B-5
16	49	25.9	488	3	US-09-069-781B-5
17	49	25.9	488	4	US-09-337-132-5
18	49	25.9	488	4	US-08-864-564A-5
19	49	25.9	488	4	US-09-094-410-5
20	49	25.9	488	4	US-08-708-123D-5
21	49	25.9	488	4	US-08-583-153A-5
22	49	25.9	488	4	US-08-638-524B-5
23	49	25.9	431	4	US-09-107-532A-6380
24	49	25.9	572	2	US-08-419-652-5
25	49	25.9	655	4	US-09-632-538C-36
26	49	25.9	658	2	US-08-825-558-4
27	49	25.9	658	4	US-09-312-611-4

28	49	25.9	708	1	US-07-797-556-2	Sequence 2, Appli
29	49	25.9	708	1	US-08-308-881-2	Sequence 2, Appli
30	49	25.9	708	2	US-09-058-263-2	Sequence 2, Appli
31	49	25.9	708	2	US-09-059-099-2	Sequence 2, Appli
32	49	25.9	708	3	US-09-058-264-2	Sequence 2, Appli
33	49	25.9	708	4	US-09-455-962-2	Sequence 2, Appli
34	49	25.9	708	5	PCT-US95-06530-2	Sequence 2, Appli
35	49	25.9	859	4	US-09-313-942-7	Sequence 7, Appli
36	49	25.9	918	2	US-08-825-558-6	Sequence 6, Appli
37	49	25.9	918	4	US-09-312-611-6	Sequence 6, Appli
38	49	25.9	951	4	US-09-313-942-9	Sequence 9, Appli
39	49	25.9	1002	4	US-09-762-724-4	Sequence 4, Appli
40	49	25.9	1158	4	US-09-313-942-26	Sequence 26, Appli
41	49	25.9	1168	4	US-09-313-942-24	Sequence 24, Appli
42	48.5	25.7	1428	1	US-07-882-790-4	Sequence 4, Appli
43	48.5	25.7	1129	4	US-09-543-681A-6416	Sequence 6416, Ap
44	48	25.4	371	4	US-09-653-375B-9	Sequence 9, Appli
45	48	25.4	803	4	US-09-540-236-2588	Sequence 2588, Ap

ALIGNMENTS

RESULT 1
US-09-107-532A-4726
; Sequence 4726, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4726:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 623 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...623
; SEQUENCE DESCRIPTION: SEQ ID NO: 4726:
US-09-107-532A-4726

DE ORF27.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=569b;
RX MEDLINE=20476433; PubMed=11021935;
RA Clark C.A., Putins L., Kaewrakon P., Focareta T., Manning P.A.;
RT "The Vibrio cholerae O1 chromosomal integron."
RL Microbiology 146:2605-2612(2000).
DR EMBL; AFI79596; AAF71189.1; -.
SQ SEQUENCE 301 AA; 33413 MW; F0D66AF91E11609F CRC64;

Query Match 40.0%; Score 44; DB 2; Length 301;
Best Local Similarity 47.4%; Pred.No. 48;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 ARKKTFLSVHEVMAVENYA 20
|||:|:|:|:
Db 208 ARVNGFSSIHMMGTVEFA 226

Search completed: March 18, 2004, 14:21:25
Job time : 15.1825 secs

THIS PAGE BLANK (USPIC)

```

DR EMBL; AP005144; BAC64255.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003678; F:DNA helicase activity; IEA.
DR GO; GO:0006260; F:DNA replication; IEA.
DR InterPro; IPR007694; DnaB_C.
DR InterPro; IPR007693; DnaB_N.
DR Pfam; PF00772; DnaB; 1.
DR PFam; PF03796; DnaB C; 1.
SQ SEQUENCE 451 AA; 50587 MW; D0EAA985B134390F CRC64;

Query Match 40.5%; Score 44.5; DB 16; Length 451;
Best Local Similarity 45.0%; Pred. No. 61;
Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 2 ARKTFLSVHEVMAVENYAK 21
: : : : : : : : : : : : : : : :
DB 156 SNKSFRIHDVLA-ENYQK 174

RESULT 12
Q8IC13 Q8IC13 PRELIMINARY; PRT; 1025 AA.
AC Q8IC13;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN MAL6P1.239.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,
RA Berriman M., Pain N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844505; CAP50459.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1025 AA; 123081 MW; 6AA4751E25D5F7FF CRC64;

Query Match 40.5%; Score 44.5; DB 5; Length 1025;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 8; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 3 RKKTFLSVH-EVMAVENY 19
: : : : : : : : : : : : : : : :
DB 697 QKSFMSIHDIYNILNY 714

RESULT 13
Q8IJT8 Q8IJT8 PRELIMINARY; PRT; 175 AA.
AC Q8IJT8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA polymerase sigma-70 factor, ECF subfamily.
GN BA5610.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,

RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017041; AAP29248.1; -.
DR TIGR; BA5610; -.
DR InterPro; IPR009043; RNA_pol_sigma.
DR InterPro; IPR007627; Sigma70_r2.
DR Pfam; PF04542; sigma70_r2; 1.
DR Pfam; PF04545; sigma70_r4; 1.
KW Complete proteome.
SQ SEQUENCE 175 AA; 21395 MW; 58583126E6657D80 CRC64;

Query Match 40.0%; Score 44; DB 16; Length 175;
Best Local Similarity 52.9%; Pred. No. 27;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 KTFSLVHEVMAVENYAK 21
: : : : : : : : : : : : : : : :
DB 40 KVFSSRHQLSYVENYKK 56

RESULT 14
Q814Q9 Q814Q9 PRELIMINARY; PRT; 176 AA.
AC Q814Q9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA polymerase ECF-type sigma factor.
GN BC5383.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapural V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goitsman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis.";
RL Nature 423:87-91(2003).
DR EMBL; AE017015; AAP12225.1; -.
DR InterPro; IPR009043; RNA_pol_sigma.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007630; Sigma70_r4.
DR Pfam; PF04542; sigma70_r2; 1.
DR Pfam; PF04545; sigma70_r4; 1.
KW Complete proteome.
SQ SEQUENCE 176 AA; 21582 MW; 46285FF1A4B528A1 CRC64;

Query Match 40.0%; Score 44; DB 16; Length 176;
Best Local Similarity 52.9%; Pred. No. 28;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 KTFSLVHEVMAVENYAK 21
: : : : : : : : : : : : : : : :
DB 41 KVFSSRHQLSYVENYKK 57

RESULT 15
Q9L985 Q9L985 PRELIMINARY; PRT; 301 AA.
AC Q9L985;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```



```

DR InterPro; IPR002550; CBS.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF01595; DUF21; 1.
SQ SEQUENCE 706 AA; 77711 MW; AAB3EACAD8E5F53 CRC64;

Query Match      40.9%; Score 45; DB 3; Length 706;
Best Local Similarity 44.4%; Pred. No. 80;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      5 KTFSLVHEVMVAVENYAKD 22
Db      222 KTLVTLHRTMGVERLTKD 239

RESULT 9
QNG54      PRELIMINARY; PRT; 852 AA.
AC QNG54;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aeryl hydrocarbon receptor-like protein.
OS Mya arenaria.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida;
OC Myoidea; Myidae; Mya.
OX NCBI_TaxID=5604;
RN [1]
RP SEQUENCE FROM N.A.
RA Kelley M.L., Powell W.H., Hahn M.E., Van Beneden R.J.;
RT "An Aeryl Hydrocarbon Receptor from the Softshell Clam, Mya arenaria.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
    TRANSCRIPTION FACTORS.
DR EMBL; AF261769; AAF70378.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC_domain.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 1.
DR PROSITE; PS00888; HLH_2; 1.
DR PROSITE; PS00112; PAS; 1.
KW Receptor.
SQ SEQUENCE 852 AA; 96580 MW; 06353220C22ABC10 CRC64;

Query Match      40.9%; Score 45; DB 5; Length 852;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      5 KTFSLVHEVMVAVENY 19
Db      186 KANLTLHEVMVMPENY 200

RESULT 10
Q8P213      PRELIMINARY; PRT; 446 AA.
AC Q8P213;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative DNA polymerase III delta prime subunit (Putative DNA
DE polymerase III delta prime subunit-phage associated).
GN SPY18_0348 OR SPYN3_0690.
OS Streptococcus pyogenes (serotype M18), and
OS Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=1212206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR EMBL; AS009979; AAL97100.1; -.
DR EMBL; AE014150; AAM79297.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003678; F:DNA helicase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR007694; DnaB_C.
DR InterPro; IPR007692; DnaB_helicase.
DR InterPro; IPR007693; DnaB_N.
DR Pfam; PF00772; DnaB; 1.
DR Pfam; PF03796; DnaB_C; 1.
DR ProDom; PD332834; DnaB_C; 1.
DR TIGRFAMs; TIGR00665; DnaB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 446 AA; 50028 MW; FD98A785AAA3BE2E CRC64;

Query Match      40.5%; Score 44.5; DB 16; Length 446;
Best Local Similarity 45.0%; Pred. No. 60;
Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY      2 ARKTFSLVHEVMVAVENYAK 21
Db      151 SNKSPRPIDVLA-ENYOK 169

RESULT 11
Q878M1      PRELIMINARY; PRT; 451 AA.
AC Q878M1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative DNA polymerase III delta prime subunit.
GN SPS1160.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayaashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```

```
Query Match      42.7%; Score 47; DB 16; Length 105;
Best Local Similarity 31.8%; Pred.No. 4.9;
Matches 7; Conservative 11; Mismatches 2; Indels 2; Gaps 1;
```

```

RESULT 8
Q12296 PRELIMINARY; PRT; 706 AA.
AC Q12296;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Chromosome XV reading frame ORF YOL060C.
GN MAM3 OR AMI3 OR YOL060C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID:4932;
[1]
RN SEQUENCE FROM N.A.
RA Feldmann H., Mannhaupt G., Vetter I.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=ALPHA S288;
RA Mannhaupt G., Vetter I., Schwarzlose C., Mitzel S., Feldmann H.;
RL Yeast 0-0-0(0).
DR EMBL; Z74802; CAA99069.1; -.
DR EMBL; X91067; CAA62524.1; -.
DR PIR; S61717; S61717.
DR SGD; S0005421; MAM3.
DR GO; GO:0007005; P.mitochondrion organization and biogenesis; IMP;

```

```
DE SC13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335480; AAL71991.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 138 AA; 16506 MW; B49440ACA3585C64 CRC64;

Query Match 100.0%; Score 110; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QARKKTFLSVHEVMAVENYAKD 22
DB 25 QARKKTFLSVHEVMAVENYAKD 46

RESULT 3
O8K5A3 PRELIMINARY; PRT; 139 AA.
ID Q8K5A3
AC Q8K5A3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin 11.
GN CST11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Hall S.H.;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF501290; AAM21709.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 139 AA; 16686 MW; B1E36DB786B4D08C CRC64;

Query Match 48.2%; Score 53; DB 11; Length 139;
Best Local Similarity 45.5%; Pred. No. 0.63;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 QARKKTFLSVHEVMAVENYAKD 22
DB 25 QVKKRTFIRVEVNALESSVKE 76

RESULT 4
Q927X9 PRELIMINARY; PRT; 184 AA.
ID Q927X9
AC Q927X9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein CPn0574.
GN CPN0574 OR CPJ0574 OR CP0175 OR CPB0596.
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=9920606; PubMed=10192388;
RA Kalnan S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Uettermann T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871162;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS001642; AAD18714.1; -.
DR EMBL; AR002178; AAF38049.1; -.
DR EMBL; AP002547; BAA98780.1; -.
DR EMBL; AE017159; AAP98525.1; -.
DR PIR; B8562; B8562.
DR PIR; D72061; D72061.
DR TIGR; CP0175; -.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 184 AA; 22024 MW; EF36820D1A97A814 CRC64;

Query Match 48.2%; Score 53; DB 16; Length 184;
Best Local Similarity 55.0%; Pred. No. 0.84;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
DB 26 RKKTFLSHRVLARPSFPVD 45

RESULT 5
Q9BIA2 PRELIMINARY; PRT; 1227 AA.
ID Q9BIA2
AC Q9BIA2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C33F10.5 OR C33F10.5B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 13.1825 Seconds
(without alignments)

526.560 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFSLVHEVMAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	103	4 Q8WXU5	Q8WXU5 homo sapien
2	110	100.0	138	4 Q8WXU6	Q8WXU6 homo sapien
3	53	48.2	139	11 Q8K5A3	Q8K5A3 rattus norv
4	53	48.2	184	16 Q9Z7X9	Q9Z7X9 chlamydia p
5	48	43.6	1227	5 Q9BIA2	Q9BIA2 caenorhabdi
6	47	42.7	105	16 Q8DB15	Q8DB15 vibrio vuln
7	45.5	41.4	273	17 Q8ZZ45	Q8ZZ45 pyrobaculum
8	45	40.9	706	3 Q12296	Q12296 saccharomyc
9	45	40.9	852	5 Q9NG54	Q9NG54 mya arenari
10	44.5	40.5	446	16 Q8P2I3	Q8P2I3 streptococc
11	44.5	40.5	451	16 Q878M1	Q878M1 streptococc
12	44.5	40.5	1025	5 Q8IC13	Q8IC13 plasmodium
13	44	40.0	175	16 Q81J78	Q81J78 bacillus an
14	44	40.0	176	16 Q814Q9	Q814Q9 bacillus ce
15	44	40.0	301	2 Q9L985	Q9L985 vibrio chol
16	44	40.0	542	4 Q96M62	Q96M62 homo sapien

17	44	40.0	576	4 Q81Y28	Q81Y28 homo sapien
18	44	40.0	704	4 Q96EF0	Q96EF0 homo sapien
19	43	39.1	102	10 Q49991	Q49991 lycopersico
20	43	39.1	127	9 Q8SCR2	Q8SCR2 pseudomonas
21	43	39.1	212	5 Q966N6	Q966N6 caenorhabdi
22	43	39.1	259	16 Q25194	Q25194 helicobacte
23	43	39.1	294	16 Q88BQ6	Q88BQ6 pseudomonas
24	43	39.1	366	5 Q93542	Q93542 caenorhabdi
25	43	39.1	394	2 Q8RK17	Q8RK17 pseudomonas
26	43	39.1	428	16 Q88C18	Q88C18 pseudomonas
27	43	39.1	431	16 Q9JXY5	Q9JXY5 neisseria m
28	43	39.1	431	16 Q9JYV6	Q9JYV6 neisseria m
29	43	39.1	467	5 Q19003	Q19003 caenorhabdi
30	43	39.1	488	10 Q9SEJ7	Q9SEJ7 lupinus alb
31	43	39.1	493	10 Q94E49	Q94E49 oryza sativ
32	43	39.1	585	16 Q97KY8	Q97KY8 clostridium
33	43	39.1	676	10 Q48785	Q48785 arabidopsis
34	43	39.1	1217	16 Q82241	Q82241 enterococcu
35	42	38.2	117	15 Q9VYW4	Q9VYW4 human immun
36	42	38.2	119	10 Q9S9D5	Q9S9D5 glycine max
37	42	38.2	122	15 Q90DQ6	Q90DQ6 human immun
38	42	38.2	127	15 Q9YZ16	Q9YZ16 human immun
39	42	38.2	140	5 Q8T5B0	Q8T5B0 plasmodium
40	42	38.2	154	16 Q8XI64	Q8XI64 clostridium
41	42	38.2	208	2 Q9F6B5	Q9F6B5 pseudoalter
42	42	38.2	209	2 Q68663	Q68663 neptunomona
43	42	38.2	254	10 Q84RU4	Q84RU4 oryza sativ
44	42	38.2	256	16 Q99RZ1	Q99RZ1 staphylococ
45	42	38.2	256	16 Q8NV94	Q8NV94 staphylococ

ALIGNMENTS

RESULT 1

Q8WXU5 PRELIMINARY; PRT; 103 AA.

AC Q8WXU5
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE SC13delta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335481; AAL71992.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 100.0%; Score 110; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 8.8e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QARKKTFSLVHEVMAVENYAKD 22

Db 25 QARKKTFSLVHEVMAVENYAKD 46

RESULT 2

Q8WXU6 PRELIMINARY; PRT; 138 AA.

AC Q8WXU6
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

Db 28 AKKSNFLEPLAAHSYLAVEFYA 49

Search completed: March 18, 2004, 14:16:09
Job time : 3.61905 secs


```

RR4 RICNA
ID RR4 RICNA STANDARD; PRT; 202 AA.
AC Q9M4C2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4.
GN RPS4.
OS Ricciocarpos natans (Liverwort).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiopsida; Marchantiidae; Ricciales; Ricciaceae; Ricciocarpos.
OX NCBI_TaxID=53035;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gametophyte;
RA Capsius I.; Blocher R.;
RT "A molecular approach to bryophyte systematics.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -!- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -!- SIMILARITY: Belongs to the S4P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ251062; CAB92155.1; -
CC HSSP; P81288; IC05. -
CC HMAP; MF_01306; -; 1.
CC InterPro; IPR001912; Ribosomal_S4.
CC InterPro; IPR005709; Ribosomal_S4_b/o.
CC Pfam; PF00163; Ribosomal_S4; 1.
CC Pfam; PF01479; S4; 1.
CC SMART; SM00363; S4; 1.
CC TIGRFAMs; TIGR01017; rpsd_bact; 1.
CC PROSITE; PS00632; RIBOSOMAL_S4; 1.
CC PROSITE; PS50889; S4; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
PT DOMAIN 90 154 S4 RNA-BINDING.
SQ SEQUENCE 202 AA; 23565 MW; 607EBC24ED641FC1 CRC64;

Query Match 37.3%; Score 41; DB 1; Length 202;
Best Local Similarity 36.8%; Pred. No. 12;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAK 21
DB 182 REWIFLKNELLVVEYYSR 200
|: || :||: || :|
|: || :||: || :|

RESULT 13
GSHR_HAEIN STANDARD; PRT; 456 AA.
AC P43783;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glutathione reductase (EC 1.8.1.7) (GR) (Grase).
GN GOR OR H10161.

```

```

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RA Barcak G.J.; Heimer S.R.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=9550630; PubMed=7542800;
RA Fleischmann R.D.; Adams M.D.; White O.; Clayton R.A.; Kirkness E.F.;
RA Kerlavage A.R.; Bult C.J.; Tomb J.-P.; Dougherty B.A.; Merrick J.M.;
RA McInerney K.; Sutton G.; Fitzhugh W.; Fields C.A.; Gocayne J.D.;
RA Scott J.D.; Shirley R.; Liu L.-I.; Glodek A.; Kelley J.M.;
RA Weidman J.P.; Phillips C.A.; Spriggs T.; Hedblom E.; Cotton M.D.;
RA Uterback T.R.; Hanna M.C.; Nguyen D.T.; Saudek D.M.; Brandon R.C.;
RA Fine L.D.; Fritchman J.L.; Fuhrmann J.L.; Geoghagen N.S.M.;
RA Gnehm C.L.; McDonald L.A.; Small K.V.; Fraser C.M.; Smith H.O.;
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: Maintain high levels of reduced glutathione in the
CC cytosol (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 glutathione + NADP(+) = glutathione
CC disulfide + NADPH.
CC -!- COFACTOR: Binds 1 FAD per subunit (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: The active site is a redox-active disulfide bond.
CC -!- SIMILARITY: Belongs to class-I pyridine nucleotide-disulfide
CC oxidoreductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32029; AAA62137.1; -
CC EMBL; U32702; AAC21833.1; -
CC PIR; A64052; A64052.
CC HSSP; P00350; IALG.
CC TIGR; H10161; -
CC InterPro; IPR001327; FAD_pyr_redox.
CC InterPro; IPR006322; Glut_reduct_1.
CC InterPro; IPR000815; Hg_reductase.
CC InterPro; IPR001100; Pyr_redox.
CC InterPro; IPR004099; Pyr_redox_dim.
CC InterPro; IPR000103; Pyridine_redox_2.
CC Pfam; PF00070; pyr_redox; 1.
CC Pfam; PF02852; pyr_redox_dim; 1.
CC PRINTS; PR00368; FADPNR.
CC PRINTS; PR00945; HGRDTASE.
CC PRINTS; PR00411; PNDRTASEI.
CC PRINTS; PR00469; PNDRTASEII.
CC PRODOM; PD000139; FAD_pyr_redox; 1.
CC TIGRFAMs; TIGR01421; Gluta_reduc_1; 1.
CC PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP;
KW Complete proteome.
FT NP_BIND 34 42 FAD (ADP PART) (BY SIMILARITY).
FT DISULFID 42 47 REDOX-ACTIVE (BY SIMILARITY).
FT ACT_SITE 445 445 BY SIMILARITY.
SQ SEQUENCE 456 AA; 49329 MW; 84698C215FE949A CRC64;

Query Match 37.3%; Score 41; DB 1; Length 456;
Best Local Similarity 40.0%; Pred. No. 30;

```

DR InterPro; IPR008994; Nucleic acid OB.
 DR InterPro; IPR004364; tRNA-synt_2.
 DR InterPro; IPR002332; tRNA-synt_2.
 DR InterPro; IPR004365; tRNA anti.
 DR InterPro; IPR006195; tRNA ligase II.
 DR Pfam; PF00152; tRNA-synt_2; 1.
 DR Pfam; PF01336; tRNA anti; 1.
 DR PRINTS; PR01042; TENASYNTHASP.
 DR TIGRFAMs; TIGR00457; aenS; 1.
 DR PROSITE; PSS0862; AA TRNA LIGASE II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 466 AA; 54136 MW; 52946E567AD0E826 CRC64;
 Query Match 38.2%; Score 42; DB 1; Length 466;
 Best Local Similarity 47.1%; Pred. No. 21;
 Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 4 KKTFLSVHEVMAVENVA 20
 DB 204 KRTFLVSGQLHIESYA 220
 RESULT 10
 LON1 THEAC STANDARD; PRT; 657 AA.
 AC Q9HJ89; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative protease La homolog type 1 (EC 3.4.21.-).
 GN TA1081.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasmata.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
 RT acidophilum";
 RL Nature 407:508-513(2000).
 CC -1- SIMILARITY: Belongs to peptidase family S16.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AL445066; CAC12209.1; --
 DR MEROPS; S16.005; --
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR004663; Pept S16 Archaea.
 DR InterPro; IPR008269; Pept S16 C.
 DR InterPro; IPR001984; Peptidase_S16.
 DR Pfam; PF05362; Lon_C; 1.
 DR PRINTS; PR00830; ENDOLAPTASE.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00764; Lon rel; 1.
 KW Hypothetical protein; Hydrolase; Serine protease; ATP-binding;
 KW Complete proteome.
 FT NP_BIND 57 64 ATP (POTENTIAL).
 FT ACT_SITE 525 525 BY SIMILARITY.
 SQ SEQUENCE 657 AA; 71600 MW; FD465CD99F5B38F7 CRC64;
 Query Match 38.2%; Score 42; DB 1; Length 657;
 Best Local Similarity 40.0%; Pred. No. 30;

Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 QY 2 ARKKTFLSVHEVMAVENYAK 21
 DB 390 SQKKTVTVTAAADVIAAKNLAK 409
 RESULT 11
 RR4 HOOLU STANDARD; PRT; 202 AA.
 ID AC P59139;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S4.
 GN RPS4.
 OS Hookeria lucens (Moss).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 OC Bryopsida; Bryidae; Hookeriales; Hookeriaceae; Hookeria.
 OX NCBI_TaxID=65539;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Gametophyte;
 RA Blocher R., Capesius I.;
 RT "The systematic position of the Hypopterygiaceae (Bryopsida) inferred
 RT from rps4 gene sequences.";
 RL Cryptogam. Bryol. 23:191-207(2002).
 CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -1- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
 CC -1- SIMILARITY: Belongs to the S4p family of ribosomal proteins.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AJ269689; CAC80629.1; --
 DR HAMAP; MF_01306; 1.
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR005709; Ribosomal_S4_b/o.
 DR InterPro; IPR002942; S4.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsD_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE; PS00889; S4; 1.
 KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
 FT DOMAIN 90 153 S4 RNA-BINDING.
 SQ SEQUENCE 202 AA; 23506 MW; C77BDB4670B62D5A CRC64;
 Query Match 37.3%; Score 41; DB 1; Length 202;
 Best Local Similarity 36.8%; Pred. No. 12;
 Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 QY 3 RKKTFLSVHEVMAVENYAK 21
 DB 182 RESTGLKINELLYVEYYSR 200
 RESULT 12


```

FT DOMAIN 517 553 EGF-LIKE.
FT TSP TYPE-1 3.
FT DOMAIN 565 612 TSP-BINDING DOMAIN.
FT DOMAIN 642 934 C5B-BINDING DOMAIN.
FT DOMAIN 643 700 SUSHI 1.
FT DOMAIN 703 762 SUSHI 2.
FT DOMAIN 766 840 COMPLEMENT CONTROL FACTOR I MODULE 1.
FT DOMAIN 858 934 COMPLEMENT CONTROL FACTOR I MODULE 2.
FT DISULFID 140 151 BY SIMILARITY.
FT DISULFID 146 164 BY SIMILARITY.
FT DISULFID 158 173 BY SIMILARITY.
FT DISULFID 399 420 BY SIMILARITY.
FT DISULFID 523 539 BY SIMILARITY.
FT DISULFID 541 541 BY SIMILARITY.
FT DISULFID 543 552 BY SIMILARITY.
FT DISULFID 644 686 BY SIMILARITY.
FT DISULFID 672 699 BY SIMILARITY.
FT DISULFID 704 746 BY SIMILARITY.
FT DISULFID 732 761 BY SIMILARITY.
FT CARBOHYD 29 32 C-LINKED (MAN).
FT CARBOHYD 32 32 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 90 90 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 568 568 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 571 571 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 574 574 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 119 119 E -> A (in allotype C6 B).
FT SEQUENCE 934 AA; 5A2713AB5B507B80 CRC64;

Query Match 38.6%; Score 42.5; DB 1; Length 934;
Best Local Similarity 39.1%; Pred. No. 36;
Matches 9; Conservative 6; Mismatches 5; Indels 3; Gaps 1;

QY 3 RKTLSVHEVMAVENY---AKD 22
DB 309 KDSFIRHKVMKVLNPTTKAKD 331

RESULT 8
K125 ARATH
ID K125_ARATH STANDARD; PRT; 1056 AA.
AC P82266;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 125 kDa kinesin-related protein.
GN A72G36200 OR F2H17.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronging C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
CC -1- FUNCTION: Responsible for microtubule translocation. May be
CC important for the organization of phragmoplast-specific arrays of
CC microtubules (By similarity).
CC -1- SUBCELLULAR LOCATION: Microtubule-associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. B1MC

```

```

CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AC006921; AAD21445.1; -.
CC PIR; H84777; H84777.
CC HSSP; P17119; 3KAR.
CC InterPro; IPR001752; kinesin_motor.
CC Pfam; PF00225; kinesin; 1.
CC PRINTS; PR00380; KINESINHEAVY.
CC SMART; SM00129; KISC; 1.
CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
CC PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
CC Motor protein; Microtubule; ATP-binding; Coiled coil.
FT DOMAIN 1 363 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 406 526 COILED COIL (POTENTIAL).
FT NP BIND 98 105 ATP (BY SIMILARITY).
FT SEQUENCE 1056 AA; 119267 MW; 7AD18A3FA9AB9211 CRC64;

Query Match 38.6%; Score 42.5; DB 1; Length 1056;
Best Local Similarity 61.1%; Pred. No. 41;
Matches 11; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 5 KTFLSVHEVMAVENYAKD 22
DB 770 KTFLDEH-VSAVNNLTKD 786

RESULT 9
SYN_WIGBR
ID SYN_WIGBR STANDARD; PRT; 466 AA.
AC Q8D2U1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Asparaginyl-tRNA synthetase (EC 6.1.1.22) (Asparagine--tRNA ligase)
DE (AsnRS).
GN ASNS OR WIGBR2630.
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22297718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia.";
RL Nat. Genet. 32:402-407(2002).
CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +
CC diphosphate + L-asparaginyl-tRNA(Asn).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB063521; BAC24409.1; -.
CC HAMAP; MF_00534; -, 1.
CC InterPro; IPR004522; AsnS.

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Complement component C6 precursor.
GN C6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-31 AND 633-640.
RX MEDLINE=90036879; PubMed=2808363;
RA Haeffliger J.-A., Teichopp J., Vial N., Jenne D.E.;
RT "Complete primary structure and functional characterization of the
RT sixth component of the human complement system. Identification of the
RT C5b-binding domain in complement C6.";
RL J. Biol. Chem. 264:18041-18051(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89380223; PubMed=2789218;
RA Discipio R.G., Hugli T.E.;
RT "The molecular architecture of human complement component C6.";
RL J. Biol. Chem. 264:16197-16206(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93291175; PubMed=8512929;
RA Hobart M.J., Fernie B., Discipio R.G.;
RT "Structure of the human C6 gene.";
RL Biochemistry 32:6198-6205(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenner C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 1-491 FROM N.A.
RX MEDLINE=89202413; PubMed=2468158;
RA Chakravarti D.N., Chakravarti B., Parra C.A., Mueller-Eberhard H.J.;
RT "Structural homology of complement protein C6 with other
RT channel-forming proteins of complement.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2799-2803(1989).
RN [6]
RP VARIANT ALLOTYPES C6 A AND C6 B.
RX MEDLINE=93326158; PubMed=8101442;
RA Dewald G., Nothen M.M., Cichon S.;
RT "Polymorphism of human complement component C6: an amino acid
RT substitution (Glu/Ala) within the second thrombospondin repeat
RT differentiates between the two common allotypes C6 A and C6 B.";
RL Biochem. Biophys. Res. Commun. 194:458-464(1993).
RN [7]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=20020247; PubMed=10551839;
RA Hofsteenge J., Blommestein M., Hess D., Furmanek A., Miroshnichenko O.;

RT "The four terminal components of the complement system are
RT C-mannosylated on multiple tryptophan residues.";
RL J. Biol. Chem. 274:32786-32794(1999).
CC -I- FUNCTION: Involved in the formation of the lytic c5b-9m complex.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- PFM: ALL CYSTEINE RESIDUES ARE ASSUMED TO BE CROSS-LINKED TO ONE
CC ANOTHER. INDIVIDUAL MODULES CONTAINING AN EVEN NUMBER OF CONSERVED
CC CYSTEINE RESIDUES ARE SUPPOSED TO HAVE DISULFIDE LINKAGES ONLY
CC WITHIN THE SAME MODULE.
CC -I- POLYMORPHISM: The sequence shown is that of allotype C6 A.
CC -I- SIMILARITY: Belongs to the complement C6/C7/C8/C9 family.
CC -I- SIMILARITY: Contains 2 Sushi (SCR) domains.
CC -I- SIMILARITY: Contains 3 TSP type-1 domains.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J05064; AAA51860.1; -;
DR EMBL: J05024; AAA59668.1; -;
DR EMBL: X72177; CAAS0994.1; -;
DR EMBL: BC035723; AAH35723.1; -;
DR EMBL: J04506; AAB59433.1; -;
DR FIR: A34372; A34372.
DR HSSP: P01130; 1AJJ.
DR Genew: HGNC:1339; C6.
DR MIM: 217050; -;
DR GO: GO:0003811; F:complement activity; TAS.
DR GO: GO:0006955; P:immune response; TAS.
DR GO: GO:0009618; P:response to pathogenic bacteria; TAS.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR003884; PacI_Mac.
DR InterPro: IPR002350; kazal.
DR InterPro: IPR002172; LDL_receptor_A.
DR InterPro: IPR001862; MAC_perforin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR000885; TSP_1.
DR Pfam: PF00050; kazal; 1.
DR Pfam: PF00057; ldl_recept_a; 1.
DR Pfam: PF01823; MACPF; 1.
DR Pfam: PF00084; sushi; 2.
DR Pfam: PF00090; tsp; 1; 3.
DR PRINTS: PR00764; COMPLEMENTC9.
DR PRINTS: PR01705; TSPREPEAT.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00057; FIMAC; 2.
DR SMART: SM00280; KAZAL; 1.
DR SMART: SM00192; LDLA; 1.
DR SMART: SM00457; MACPF; 1.
DR SMART: SM00209; TSPI; 3.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
DR PROSITE: PS01209; LDLRA_1; 1.
DR PROSITE: PS00668; LDLRA_2; 1.
DR PROSITE: PS00279; MAC_PERFORIN; 1.
DR PROSITE: PS0092; TSPI; 3.
KW Complement pathway; Glycoprotein; Plasma; Membrane attack complex;
KW Cytolysis; Sushi; Repeat; Signal; Polymorphism; EGF-like domain;
KW Transmembrane.
FT SIGNAL 1 21 COMPLEMENT COMPONENT C6.
FT CHAIN 22 934 TSP TYPE-1 1.
FT DOMAIN 81 134 TSP TYPE-1 2.
FT DOMAIN 138 175 LDL-RECEPTOR CLASS A.
FT TRANSMEM 331 349 POTENTIAL.
FT TRANSMEM 354 373 POTENTIAL.

```

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: Orotidine 5'-phosphate =UMP + CO(2).
CC -I- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SIMILARITY: Belongs to the OMP decarboxylase family. Subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57689; AAB02248.1; -.
DR HSSP; P25971; 1DBT.
DR HAMAP; MF 01200; -. 1.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Lyase; Decarboxylase; Pyrimidine biosynthesis.
FT NON TER 1 1
FT ACT_SITE 68 68 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 235 AA; 25093 MW; B5DB591AE084378 CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 235;
Best Local Similarity 50.0%; Pred. No. 3.7;
Matches 14; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

QY 1 QARKTKFLSV-----HEV-MAVENYAK 21
||||| : |||||
DB 58 QARKKVFDDMKLIDHTVARAVENIAK 85

RESULT 5
YLP3_PSEPU STANDARD; PRT; 394 AA.
AC P31049;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 44.7 kDa protein in lpd-3 5'region (ORF3).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RA Lorenz D., Sokatch J.R.;
RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X55704; CAA39234.1; -.
DR InterPro; IPR003333; CMAS.
DR InterPro; IPR001601; Methyltransf.
DR Pfam; PF02353; CMAS; 1.
KW Hypothetical protein.
SQ SEQUENCE 394 AA; 44656 MW; 8CD61FE9495302AF CRC64;

Query Match 39.1%; Score 43; DB 1; Length 394;
Best Local Similarity 36.8%; Pred. No. 12;
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 2 ARKTKFLSVHEVMAVENYA 20

RESULT 6
LBD2_ARATH STANDARD; PRT; 205 AA.
AC Q9LNB9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative LOB domain protein 2.
GN LBD2 OR ATG06280 OR F9P14.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Taiton L.J., Tambunga G., Toriumi M.J., Town C.D., Walker M.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
CC -I- SIMILARITY: Contains 1 LOB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC052590; AAF80225.1; -.
DR PIR; F86198; F86198. DUF260.
DR InterPro; IPR004883; DUF260.
DR Pfam; PF03195; DUF260; 1.
DR PROSITE; PS50891; LOB; 1.
KW Hypothetical protein.
FT DOMAIN 22 122 LOB.
SQ SEQUENCE 205 AA; 23285 MW; 404FD0C4FC32A342 CRC64;

Query Match 38.6%; Score 42.5; DB 1; Length 205;
Best Local Similarity 52.4%; Pred. No. 7;
Matches 11; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 2 ARK-TFSLSVHEVMAVENYAK 21
||||| : |||||
DB 46 ARKTKFQAVHKVFGVSNVQK 66

RESULT 7
CO6_HUMAN STANDARD; PRT; 934 AA.
ID CO6_HUMAN
AC P13671;

```

```

CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
-----
DR DR EMBL; D42125; BAA07705.1; -.
DR DR EMBL; AE003784; AAM68337.1; -.
DR DR EMBL; AY058652; AAL13881.1; -.
DR DR EMBL; S55977; AAB19907.1; -.
DR DR EMBL; AJ002911; CAA05746.1; -.
DR DR HSSP; P00523; ISRL.
DR DR FlyBase; FBgn0004603; Src42A.
DR DR GO; GO:0005912; C.adherens junction; IDA.
DR DR GO; GO:0004713; P.protein-tyrosine kinase activity; NAS.
DR DR GO; GO:0007154; P.cell communication; IDA.
DR DR GO; GO:0007456; P.eye morphogenesis (sensu Drosophila); IMP.
DR DR GO; GO:0007169; P.transmembrane receptor protein tyrosine kin. .; IGI.
DR DR InterPro; IPR000719; Prot_kinase.
DR DR InterPro; IPR000980; SH2.
DR DR InterPro; IPR001452; SH3.
DR DR InterPro; IPR001245; Tyr_pkinase.
DR DR InterPro; IPR008266; Tyr_pkinase_AS.
DR DR Pfam; PF00069; kinase; 1.
DR DR Pfam; PF00017; SH2; 1.
DR DR Pfam; PF00018; SH3; 1.
DR DR PRINTS; PR00401; SH2DOMAIN.
DR DR PRINTS; PR00452; SH3DOMAIN.
DR DR PRINTS; PR00109; TYRKINASE.
DR DR ProDom; PD000001; Prot_kinase; 1.
DR DR ProDom; PD000093; SH2; 1.
DR DR ProDom; PD000066; SH3; 1.
DR DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR DR PROSITE; PS00001; SH2; 1.
DR DR PROSITE; PS00002; SH3; 1.
DR DR Transferase; Tyrosine-protein kinase; Developmental protein;
KW ATP-binding; SH2 domain; SH3 domain.
FT FT DOMAIN 63 124
FT FT DOMAIN 130 222
FT FT DOMAIN 248 504
FT FT NP_BIND 254 262
FT FT BINDING 276 276
FT FT ACT_SITE 370 370
FT FT CONFLICT 65 65
FT FT CONFLICT 376 376
FT FT CONFLICT 381 386
FT FT SEQUENCE 517 AA; 59069 MW; 1EF196B4D7AE61E9 CRC64;
SQ
Query Match 40.9%; Score 45; DB 1; Length 517;
Best Local Similarity 47.6%; Pred.No. 7.1;
Matches 10; Conservative 6; Mismatches 3; Indels 2; Gaps 1;

Qy 2 ARKKTPLSVHEVAVENYAKD 22
Db 192 ARRTTFLQLQEL--VEHYSKD 210
|||:::||:||||
|::||::||:|

RESULT 4
PYRF_BARBA
ID PYRF BARBA STANDARD; PRT; 235 AA.
AC Q44843;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cratichne 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPDCase) (OMPDcase) (Fragment).
PYRF GN GN
OS Bartonella bacilliformis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=774;
RN [1]
RA UpeSlacis E., Ihler G.M.;
SEQUENCE FROM N.A.
```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL096677; CAC13170.1; -;
 DR EMBL; AL096677; CAC17423.1; -;
 DR HSSP; P01038; 1A90.
 DR Genew; HGNC:15959; CSTR11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KW Thiol protease inhibitor; Signal; Alternative splicing.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 137 CYSTATIN 11.
 FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 93 101 BY SIMILARITY.
 FT DISULFID 114 134 BY SIMILARITY.
 FT CARBOHYD 131 131 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARSPLIC 76 110 Missing (in isoform 2).
 FT FTId=VSP_001260.
 SQ SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 110; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QARKTFLSVHEVMAVENYAKD 22
 Db 24 QARKTFLSVHEVMAVENYAKD 45

RESULT 2

CS11 MOUSE
 ID CS11 MOUSE STANDARD; PRT; 139 AA.
 AC Q9D285;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin 11 precursor.
 GN CST11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Epithelium;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guncich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT Functional annotation of a full-length mouse cDNA collection.;
 RL Nature 409:685-690(2001).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AK020300; BAB32061.1; -;
 DR HSSP; P01034; 1G96.
 DR MGD; MGI:1925490; Cst11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KW Thiol protease inhibitor; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 139 CYSTATIN 11.
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 94 102 BY SIMILARITY.
 FT DISULFID 115 135 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 139 AA; 16217 MW; F228D9815PA32640 CRC64;

Query Match 48.2%; Score 53; DB 1; Length 139;
 Best Local Similarity 40.9%; Pred. No. 0.071;
 Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 1 QARKTFLSVHEVMAVENYAKD 22
 Db 25 QVKKRTFIRIEVSALSSVKE 46

RESULT 3

SR42 DROME
 ID SR42 DROME STANDARD; PRT; 517 AA.
 AC Q9V9J3; O18369; Q26297; Q94879;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tyrosine-protein kinase Src42A (EC 2.7.1.112) (Darc41).
 GN SRC42A OR SRC41 OR TK5 OR CG7873.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
 RC STAGE.
 RC STRAIN=Canton-S; TISSUE=Pupa;
 RX MEDLINE=96268448; PubMed=8692295;
 RA Takahashi P., Endo S., Kojima T., Saigo K.;
 RT "Regulation of cell-cell contacts in developing Drosophila eyes by
 RT Darc41, a new, close relative of vertebrate c-src";
 RL Genes Dev. 10:1645-1656(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agapay A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 2.61905 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFLSVHEWVAENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	53	48.2	139	CS11 MOUSE	Q9d269 mus musculu
3	45	40.9	517	SR42 DROME	Q9v913 drosophila
4	44.5	40.5	235	YPRF BARBA	Q44843 bartonella
5	43	39.1	394	YLP3 PSEPU	P11049 pseudomonas
6	42.5	38.6	205	1 LBD2 ARATH	Q9lnb9 arabidopsis
7	42.5	38.6	934	CO6 HUMAN	P13671 homo sapien
8	42.5	38.6	1056	K125 ARATH	P82266 arabidopsis
9	42	38.2	466	1 SYN WIGBR	Q8d2u1 wigglewort
10	42	38.2	657	1 LONI THEAC	Q9hj89 thermoplasm
11	41	37.3	202	RR4 HOOLU	P59139 hookeria lu
12	41	37.3	202	RR4 RICNA	Q9m4c2 riccioearpo
13	41	37.3	456	1 GSHR HABIN	P43783 haemophilus
14	40.5	36.8	264	1 SPED PSEAE	Q915x7 pseudomonas
15	40.5	36.8	508	1 GSBP CHICK	P12244 gallus gall
16	40	36.4	237	1 PR12 SOYBN	Q914c4 glycine max
17	40	36.4	288	1 LBD2 ARATH	Q9lrw1 arabidopsis
18	40	36.4	292	1 YZ51 MYCTU	P71850 mycobacteri
19	40	36.4	294	1 SUCD METJA	Q58643 methanococ
20	40	36.4	437	1 SUCB NEOPR	P53597 neocallima
21	40	36.4	496	1 Y048 UREPA	Q9pr97 ureaplasma
22	40	36.4	977	1 UB24 HUMAN	Q9upus homo sapien
23	40	36.4	6629	1 RIAB IBVBC	P27920 a replicase
24	40	36.4	6629	1 RIAB IBVBC	Q91qt2 a replicase
25	39.5	35.9	264	1 SPED PSESM	Q889x9 pseudomonas
26	39	35.5	202	1 RR4 HYPLA	P59145 hypoterygi
27	39	35.5	249	1 CORJ METJA	Q58223 methanococ
28	39	35.5	310	1 YGRJ YEAST	P53108 saccharomyc
29	39	35.5	382	1 DHGY CUCSA	P13443 cucumis sat
30	39	35.5	398	1 DXR PHOLL	Q7n8p3 photorhabdu
31	39	35.5	431	1 AROA AQUAE	Q67494 aquifex aeo
32	39	35.5	466	1 SYN XYLFA	Q9paf5 xylella fas
33	39	35.5	466	1 SYN_XYLFT	Q87a82 xylella fas

34	39	35.5	469	1	1A13 LYCES	Q42881 lycopersico
35	39	35.5	690	1	CAN9 HUMAN	O14815 homo sapien
36	39	35.5	888	1	YJH0 YEAST	P40361 saccharomyc
37	38.5	35.0	610	1	YK11 DROME	Q27324 drosophila
38	38.5	35.0	1523	1	DPOL THEFM	P74918 thermococcu
39	38	34.5	133	1	PABP SCHMA	P29498 schistosoma
40	38	34.5	165	1	LB21 ARATH	O59r18 arabidopsis
41	38	34.5	185	1	RR4 WOORA	P59151 woodwardia
42	38	34.5	202	1	RR4 MARPO	P06358 marchantia
43	38	34.5	202	1	RR4 PLARP	Q9m4c3 plagioclasma
44	38	34.5	205	1	RR4 MESVI	Q9mun0 mesostigma
45	38	34.5	270	1	CNX3 ARATH	Q39056 arabidopsis

ALIGNMENTS

RESULT 1

CS11_HUMAN	CS11_HUMAN	STANDARD;	PRT;	137 AA.
AC	Q9H112; Q9H113;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cystatin 11 precursor.			
GN	CST11 OR CST6L			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckie E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvaeslao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,			
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,			
RA	Whithead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			
RA	Rogers J.;			
RT	"The DNA sequence and comparative analysis of human chromosome 20.";			
RL	Nature 414:865-871(2001).			
CC	-1- SUBCELLULAR LOCATION: Secreted (Potential).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=1;			
CC	isoId=Q9H112-1; Sequence=Displayed;			
CC	Name=2;			
CC	isoId=Q9H112-2; Sequence=VSP_001260;			
CC	Note=No experimental confirmation available;			
CC	-1- SIMILARITY: Belongs to the cystatin family.			

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

A/Accession: H84777
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1056 <STO>
A/Cross-references: GB:AE002093; NID:g4510356; PIDN:AAD21445.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g36200
A/Map position: 2
C/Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match 38.6%; Score 42.5; DB 2; Length 1056;
Best Local Similarity 61.1%; Pred. No. 1.3e+02;
Matches 11; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 5 KTFLSVHEVMVAVENYAKD 22
Db 770 KTFLEDEH-VSAVNNLT KD 786

RESULT 15

D90026
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: D90026
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: D90026
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-256 <KUR>
A/Cross-references: GB:BA000018; PID:g13702081; PIDN:BAB43373.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: SA2076

Query Match 38.2%; Score 42; DB 2; Length 256;
Best Local Similarity 44.4%; Pred. No. 35;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 KTFLSVHEVMVAVENYAKD 22
Db 142 KYADAHEEMVREHYQKD 159

Search completed: March 18, 2004, 14:23:13
Job time : 5.23413 secs

A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 39.1%; Score 43; DB 2; Length 676;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ARKKTFLSVHEVWAVE 17
|||:|:|:|:|:|
Db 504 ARSKSFVGTHEYLAPE 519

RESULT 12
F86198
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86198
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86198
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-205 <STO>
A:Cross-references: GB:AE005172; MID:g8844133; PIDN:AAF80225.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 39.6%; Score 42.5; DB 2; Length 205;
Best Local Similarity 52.4%; Pred. No. 23;
Matches 11; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Qy 2 ARK-KTFLSVHEVWAVENYAK 21
|||:|:|:|:|:|
Db 46 ARKTKFQAVHKVFGVSNVQK 66

RESULT 13
A34372
complement C6 precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Nov-2000
C:Accession: A34372; A34235; A32109; A31894; A53072
R:Haefliger, J.A.; Tschopp, J.; Vial, N.; Jenne, D.E.
J. Biol. Chem. 264, 18041-18051, 1989
A:Title: Complete primary structure and functional characterization of the sixth component
A:Reference number: A34372; MUID:90036879; PMID:2808363
A:Accession: A34372
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-934 <HAE>
A:Cross-references: GB:J05064; NID:g179703; PIDN:AAAS1860.1; PID:g179704
R:DiScipio, R.G.; Hugli, T.E.
J. Biol. Chem. 264, 16197-16206, 1989
A:Title: The molecular architecture of human complement component C6.
A:Reference number: A34235; MUID:69380223; PMID:2789218
A:Accession: A34235
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-118 'E', 120-934 <DIS>
A:Cross-references: GB:J05024; NID:g197824; PIDN:AAAS9668.1; PID:g307228
R:Chakravarti, D.N.; Chakravarti, B.; Parra, C.A.; Muller-Eberhard, H.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 2799-2803, 1989
A:Title: Structural homology of complement protein C6 with other channel-forming protein

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 KKTFLSVHVEVMAVENYAKD 22
 Db 328 KKVKFSLHNPITIELYSKE 346

RESULT 7

A81037
 tyrosyl-tRNA synthetase NMB1835 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
 C;Species: Neisseria meningitidis
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C;Accession: A81037
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiagnani, V.; Pizzza, M.
 Science 287, 1809-1815, 2000
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, J.
 A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A;Reference number: A81000; MUID:20175755; PMID:10710307
 A;Accession: A81037
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-431 <TET>
 A;Cross-references: GB:AE002534; GB:AE002098; NID:g7227095; PIDN:AAF42170.1; PID:g7227095
 A;Experimental source: serogroup B, strain MC58
 C;Genetics:
 A;Gene: NMB1835
 C;Superfamily: tyrosine-tRNA ligase

Query Match 39.1%; Score 43; DB 2; Length 431;
 Best Local Similarity 58.8%; Pred. No. 41;
 Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 6 TFLSVHVEVMAVENYAKD 22
 Db 272 TFLSIEIDAIE--AKD 286

RESULT 8

H81981
 probable tyrosine-tRNA ligase (EC 6.1.1.1) NMA0620 [imported] - Neisseria meningitidis (strain MC58)
 C;Species: Neisseria meningitidis
 C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
 C;Accession: H81981
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, P.; Holroyd, S.; Jørgensen, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream, A.; et al.
 Nature 404, 502-506, 2000
 A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A;Reference number: A81775; MUID:20222556; PMID:10761919
 A;Accession: H81981
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-431 <PAR>
 A;Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83910.1; PID:g7379120
 A;Experimental source: serogroup A, strain Z2491
 C;Genetics:
 A;Gene: tyrS; NMA0620
 C;Superfamily: tyrosine-tRNA ligase
 C;Keywords: ligase

Query Match 39.1%; Score 43; DB 2; Length 431;
 Best Local Similarity 58.8%; Pred. No. 41;
 Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 6 TFLSVHVEVMAVENYAKD 22
 Db 272 TFLSIEIDAIE--AKD 286

RESULT 9

T15892
 hypothetical protein D2096.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C;Accession: T15892
 R;Geisel, C.
 submitted to the EMBL Data Library, November 1995
 A;Description: The sequence of C. elegans cosmid D2096.
 A;Reference number: Z18425
 A;Accession: T15892
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-467 <GEI>
 A;Cross-references: EMBL:U40800; NID:g1065946; PID:g1065948; PIDN:AAA81490.1; CESP:D2096
 C;Genetics:
 A;Gene: CESP:D2096.4
 A;Introns: 31/1; 70/2; 122/2; 239/2; 336/3

Query Match 39.1%; Score 43; DB 2; Length 467;
 Best Local Similarity 53.3%; Pred. No. 45;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFLSVHVEVMAVENY 19
 Db 155 KLMLDGGHEVIALDNY 169

RESULT 10

G96995
 ATP-dependent RNA helicase, superfamily II [imported] - Clostridium acetobutylicum
 C;Species: Clostridium acetobutylicum
 C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C;Accession: G96995
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A;Reference number: A96900; MUID:21359325; PMID:21359325
 A;Accession: G96995
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-585 <KOR>
 A;Cross-references: GB:AE001437; PIDN:AAK78754.1; PID:g15023664; GSPDB:GN00168
 A;Experimental source: Clostridium acetobutylicum ATCC824
 C;Genetics:
 A;Gene: CAC0778

Query Match 39.1%; Score 43; DB 2; Length 585;
 Best Local Similarity 40.0%; Pred. No. 56;
 Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 ARKKTFLSVHVEVMAVENYAK 21
 Db 166 ARRGVYLSPLRILALENFEK 185

RESULT 11

G84663
 hypothetical protein At2g26700 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cross)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 19-Jul-2002
 C;Accession: G84663
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Nishizawa, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; et al.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: G84663
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-676 <STO>
 A;Cross-references: GB:AE002093; NID:g2760836; PIDN:AAB95304.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g26700

Db 105 KALLSVHDTVAKNY 119

RESULT 13

US-09-783-320-22
 ; Sequence 22, Application US/09783320
 ; Patent No. US20020038011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Turner, C. Alexander Jr
 ; APPLICANT: Zambrowicz, Brian
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0137-USA
 ; CURRENT APPLICATION NUMBER: US/09/783,320
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 60/183,582
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/184,014
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 22
 ; LENGTH: 473
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-783-320-22

Query Match 40.0%; Score 44; DB 9; Length 473;
 Best Local Similarity 53.3%; Pred. No. 84;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 KTFSLVHEVMAVENY 19
 | ||||: :||
 Db 105 KALLSVHDTVAKNY 119

RESULT 14

US-09-783-320-30
 ; Sequence 30, Application US/09783320
 ; Patent No. US20020038011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Turner, C. Alexander Jr
 ; APPLICANT: Zambrowicz, Brian
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0137-USA
 ; CURRENT APPLICATION NUMBER: US/09/783,320
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 60/183,582
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/184,014
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 30
 ; LENGTH: 520
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-783-320-30

Query Match 40.0%; Score 44; DB 9; Length 520;
 Best Local Similarity 53.3%; Pred. No. 94;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 KTFSLVHEVMAVENY 19
 | ||||: :||
 Db 105 KALLSVHDTVAKNY 119

RESULT 15

US-10-094-749-3093
 ; Sequence 3093, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHIKO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOTYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 3381
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3093
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-094-749-3093

Query Match 40.0%; Score 44; DB 15; Length 542;
 Best Local Similarity 53.3%; Pred. No. 98;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 KTFSLVHEVMAVENY 19
 | ||||: :||
 Db 105 KALLSVHDTVAKNY 119

Search completed: March 18, 2004, 14:30:52
 Job time : 13.3135 secs

SOFTWARE: SeqWin99, version 1.02
 SEQ ID NO 375
 LENGTH: 184
 TYPE: PRT
 ORGANISM: Chlamydia pneumoniae
 US-10-312-273-375

Query Match 48.2%; Score 53; DB 15; Length 184;
 Best Local Similarity 55.0%; Pred. No. 0.91;
 Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
 Db 26 RKKTFLSSHRVLARPSPFVD 45

RESULT 9
 US-10-289-762-616
 ; Sequence 616, Application US/10289762
 ; Publication No. US20040006219A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffais, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/10/289,762
 ; PRIOR FILING DATE: 2003-03-27
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 616
 ; LENGTH: 195
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-289-762-616

Query Match 44.5%; Score 49; DB 15; Length 195;
 Best Local Similarity 50.0%; Pred. No. 4.5;
 Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
 Db 37 RQKTFLSSHRVLARPSPFVD 56

RESULT 10
 US-09-783-320-8
 ; Sequence 8, Application US/09783320
 ; Patent No. US20020038011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Turner, C. Alexander Jr
 ; APPLICANT: Zambrowicz, Brian
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0137-USA
 ; CURRENT APPLICATION NUMBER: US/09/783,320
 ; CURRENT FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 60/183,582
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/184,014
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 296
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-783-320-8

Query Match 40.08; Score 44; DB 9; Length 296;
 Best Local Similarity 53.3%; Pred. No. 49;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFLSVHEVMAVENY 19
 Db 105 KALLSVHDTVAQKNY 119

RESULT 11
 US-09-783-320-12
 ; Sequence 12, Application US/09783320
 ; Patent No. US20020038011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Turner, C. Alexander Jr
 ; APPLICANT: Zambrowicz, Brian
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0137-USA
 ; CURRENT APPLICATION NUMBER: US/09/783,320
 ; CURRENT FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 60/183,582
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/184,014
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 318
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-783-320-12

Query Match 40.0%; Score 44; DB 9; Length 318;
 Best Local Similarity 53.3%; Pred. No. 53;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFLSVHEVMAVENY 19
 Db 105 KALLSVHDTVAQKNY 119

RESULT 12
 US-09-783-320-18
 ; Sequence 18, Application US/09783320
 ; Patent No. US20020038011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Turner, C. Alexander Jr
 ; APPLICANT: Zambrowicz, Brian
 ; TITLE OF INVENTION: No. US20020038011A1 Human Kinases and Polynucleotides Encoding
 ; FILE REFERENCE: LEX-0137-USA
 ; CURRENT APPLICATION NUMBER: US/09/783,320
 ; CURRENT FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 60/183,582
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/184,014
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 375
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-783-320-18

Query Match 40.0%; Score 44; DB 9; Length 375;
 Best Local Similarity 53.3%; Pred. No. 64;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KTFLSVHEVMAVENY 19
 Db 105 KALLSVHDTVAQKNY 119

; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 34822
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL109954.10
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
 ; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
 ; US-09-864-761-34822

Query Match 91.8%; Score 101; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RKKTFLSVHEVMAVENYAKD 22
 |||||
 Db 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 7
 US-09-864-761-48936
 ; Sequence 48936, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeomica-x-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 48936
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL096677.18
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
 ; OTHER INFORMATION: EST_HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
 ; US-09-864-761-48936

Query Match 91.8%; Score 101; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RKKTFLSVHEVMAVENYAKD 22
 |||||
 Db 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 8
 US-10-312-273-375
 ; Sequence 375, Application US/10312273
 ; Publication No. US20040005667A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SPA
 ; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
 ; FILE REFERENCE: P025035WO
 ; CURRENT APPLICATION NUMBER: US/10/312,273
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: 0016363.4
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 0017047.2
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 0017983.8
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: 0019368.0
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 0020440.4
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: 0022583.9
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 0027549.5
 ; PRIOR FILING DATE: 2000-11-10
 ; PRIOR APPLICATION NUMBER: 0031706.5
 ; PRIOR FILING DATE: 2000-12-22
 ; NUMBER OF SEQ ID NOS: 664

; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-3

Query Match 100.0%; Score 110; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QARKKTFSLSVHEVMAVENYAKD 22
Db 2 QARKKTFSLSVHEVMAVENYAKD 23
|||||

RESULT 3

US-09-941-314-4
; Sequence 4, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-4

Query Match 100.0%; Score 110; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QARKKTFSLSVHEVMAVENYAKD 22
Db 4 QARKKTFSLSVHEVMAVENYAKD 25
|||||

RESULT 4

US-09-941-314-2
; Sequence 2, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-2

Query Match 100.0%; Score 110; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QARKKTFSLSVHEVMAVENYAKD 22
Db 24 QARKKTFSLSVHEVMAVENYAKD 45
|||||

RESULT 5

US-09-941-314-6
; Sequence 6, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-6

Query Match 91.8%; Score 101; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFSLSVHEVMAVENYAKD 22
Db 1 RKKTFSLSVHEVMAVENYAKD 20
|||||

RESULT 6

US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 13.3135 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFSLVHEVMAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	22	9 US-09-941-314-5	Sequence 5, Appli
2	110	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
3	110	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
4	110	100.0	137	9 US-09-941-314-2	Sequence 2, Appli
5	101	91.8	36	9 US-09-941-314-6	Sequence 6, Appli
6	101	91.8	50	9 US-09-864-761-34822	Sequence 34822, A
7	101	91.8	50	9 US-09-864-761-48936	Sequence 48936, A
8	53	48.2	184	15 US-10-312-273-375	Sequence 375, App
9	49	44.5	195	15 US-10-289-762-616	Sequence 616, App
10	44	40.0	296	9 US-09-783-320-8	Sequence 8, Appli
11	44	40.0	318	9 US-09-783-320-12	Sequence 12, Appli
12	44	40.0	375	9 US-09-783-320-18	Sequence 18, Appli
13	44	40.0	473	9 US-09-783-320-22	Sequence 22, Appli
14	44	40.0	520	9 US-09-783-320-30	Sequence 30, Appli
15	44	40.0	542	15 US-10-094-749-3093	Sequence 3093, App

16	44	40.0	560	9 US-09-783-320-38	Sequence 38, Appli
17	44	40.0	576	9 US-09-783-320-46	Sequence 46, Appli
18	44	40.0	576	9 US-09-994-288-5	Sequence 5, Appli
19	44	40.0	576	16 US-10-311-034-16	Sequence 16, Appli
20	43	39.1	107	12 US-10-424-599-219730	Sequence 219730, App
21	43	39.1	259	9 US-09-895-913A-338	Sequence 338, App
22	43	39.1	431	12 US-10-282-122A-65756	Sequence 65756, A
23	43	39.1	467	15 US-10-347-470A-4	Sequence 4, Appli
24	43	39.1	1054	9 US-09-815-242-4995	Sequence 4995, App
25	43	39.1	1175	12 US-10-282-122A-57562	Sequence 57562, A
26	43	39.1	1208	9 US-09-815-242-10627	Sequence 10627, A
27	43	39.1	1217	12 US-10-282-122A-42405	Sequence 42405, A
28	42.5	38.6	281	12 US-10-282-122A-61171	Sequence 61171, A
29	42	38.2	62	12 US-10-424-599-182304	Sequence 182304, A
30	42	38.2	212	12 US-10-425-114-38286	Sequence 38286, A
31	42	38.2	273	12 US-10-424-599-227071	Sequence 227071, A
32	42	38.2	437	12 US-10-282-122A-61087	Sequence 61087, A
33	42	38.2	577	12 US-10-425-114-57760	Sequence 57760, A
34	42	38.2	597	12 US-10-424-599-207456	Sequence 207456, A
35	41	37.3	214	12 US-10-424-599-182326	Sequence 182326, A
36	41	37.3	244	12 US-10-424-599-157243	Sequence 157243, A
37	41	37.3	455	12 US-10-169-103-2	Sequence 2, Appli
38	41	37.3	456	12 US-10-282-122A-58093	Sequence 58093, A
39	41	37.3	456	14 US-10-260-877-78	Sequence 78, Appli
40	40.5	36.8	463	12 US-10-425-114-64141	Sequence 64141, A
41	40	36.4	65	12 US-10-424-599-254688	Sequence 254688, A
42	40	36.4	86	12 US-10-424-599-19963	Sequence 19963, A
43	40	36.4	102	12 US-10-424-599-177539	Sequence 177539, A
44	40	36.4	119	12 US-10-424-599-244889	Sequence 244889, A
45	40	36.4	185	12 US-10-424-599-268671	Sequence 268671, A

ALIGNMENTS

RESULT 1
US-09-941-314-5
; Sequence 5, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-5

Query Match 100.0%; Score 110; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.7e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QARKKTFSLVHEVMAVENYAKD 22
Db 1 QARKKTFSLVHEVMAVENYAKD 22

RESULT 2
US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein


```
; EARLIER APPLICATION NUMBER: 60/134,175
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Human
US-09-422-869-28
```

```
Query Match      35.5%; Score 39; DB 3; Length 690;
Best Local Similarity 53.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 QARKKTPLSVHEV 13
      :||| |||::||
Db      443 RARSKTFINLREV 455
```

```
Search completed: March 18, 2004, 14:25:35
Job time : 6.63095 secs
```


; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27435

Query Match 36.8%; Score 40.5; DB 4; Length 293;
Best Local Similarity 36.8%; Pred. No. 61;
Matches 7; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 4 KKTFLSVHEVMAVENYAKD 22
|||:|||||
Db 207 RKHFID-HEINSIQNYLSD 224

RESULT 7

US-09-800-729-195
; Sequence 195, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044PI
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 195
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-195

Query Match 36.4%; Score 40; DB 4; Length 326;
Best Local Similarity 31.2%; Pred. No. 84;
Matches 5; Conservative 7; Mismatches 4; Indels 4; Gaps 0;

QY 4 KKTFLSVHEVMAVENY 19
|||:|||||
Db 8 KKIFIHMEIIQIDGH 23

RESULT 8

US-09-489-039A-13335
; Sequence 1335, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13335
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13335

Query Match 36.4%; Score 40; DB 4; Length 419;
Best Local Similarity 47.4%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 4 KKTFLSVHEVMAVENYAKD 22
|||:|||||
Db 253 KATFSVHEAQAYETRPVD 271

RESULT 9

US-09-252-991A-25843
; Sequence 25843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25843
; LENGTH: 1242
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25843

Query Match 36.4%; Score 40; DB 4; Length 1242;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 KTFSLVHEVMAVENYAKD 22
:|||||
Db 320 ETFGVIEETTAVERLAKD 337

RESULT 10

US-09-540-236-2056
; Sequence 2056, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2056
; LENGTH: 605
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2056

Query Match 35.9%; Score 39.5; DB 4; Length 605;
Best Local Similarity 32.4%; Pred. No. 2e+02;
Matches 11; Conservative 2; Mismatches 2; Indels 19; Gaps 1;

QY 8 LSVHEVMAVE-----NYAKD 22
|||:|||||
Db 89 LSQHQVMAIEALAVYVVALFGVLLRLALIGYAKD 122

RESULT 11

US-09-328-352-6399
; Sequence 6399, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6399
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 5.63095 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKTFLSVHEVMAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	44.5	195	4	US-09-198-452A-616
2	43	39.1	255	4	US-09-543-681A-5383
3	43	39.1	1224	4	US-09-107-532A-6220
4	41	37.3	316	4	US-09-252-991A-30881
5	41	37.3	839	4	US-09-489-039A-13252
6	40.5	36.8	293	4	US-09-252-991A-27435
7	40	36.4	326	4	US-09-800-729-195
8	40	36.4	419	4	US-09-489-039A-13335
9	40	36.4	1242	4	US-09-252-991A-25843
10	39.5	35.9	605	4	US-09-540-236-2056
11	39	35.5	291	4	US-09-328-352-6399
12	39	35.5	399	4	US-09-134-000C-5712
13	39	35.5	469	3	US-08-378-313-33
14	39	35.5	664	4	US-09-377-497-70
15	39	35.5	690	3	US-09-422-869-28
16	39	35.5	1284	4	US-09-170-496B-294
17	39	35.5	1284	4	US-09-364-425B-59
18	39	35.5	1507	3	US-08-929-329-5
19	38	34.5	133	1	US-08-554-463-1
20	38	34.5	208	4	US-09-540-236-3351
21	38	34.5	271	4	US-09-489-039A-12667
22	38	34.5	284	2	US-08-956-012-3
23	38	34.5	359	4	US-09-543-681A-5431
24	38	34.5	434	4	US-09-543-681A-4577
25	38	34.5	501	1	US-08-451-715A-10
26	38	34.5	611	4	US-09-107-532A-4988
27	38	34.5	715	4	US-09-620-412C-329

28 38 34.5 715 4 US-09-598-419-329 Sequence 329, Appl
29 38 34.5 2987 2 US-08-970-269A-29 Sequence 29, Appl
30 38 34.5 2987 2 US-09-407-562-29 Sequence 29, Appl
31 38 34.5 3959 2 US-08-970-269A-30 Sequence 30, Appl
32 38 34.5 3959 3 US-09-407-562-30 Sequence 30, Appl
33 38 34.5 10182 4 US-09-134-001C-3159 Sequence 3159, Ap
34 37.5 34.1 336 4 US-09-252-991A-29131 Sequence 29131, A
35 37.5 34.1 578 1 US-08-766-014-4 Sequence 4, Appl
36 37.5 34.1 608 1 US-08-766-014-3 Sequence 3, Appl
37 37.5 34.1 872 1 US-08-766-014-2 Sequence 2, Appl
38 37 33.6 38 4 US-09-779-451-23 Sequence 23, Appl
39 37 33.6 45 4 US-09-779-451-22 Sequence 22, Appl
40 37 33.6 138 4 US-09-570-921-6 Sequence 6, Appl
41 37 33.6 190 4 US-09-540-236-3060 Sequence 3060, Ap
42 37 33.6 198 3 US-08-965-056-85 Sequence 85, Appl
43 37 33.6 198 3 US-08-965-056-90 Sequence 90, Appl
44 37 33.6 207 4 US-09-543-681A-4234 Sequence 4234, Ap
45 37 33.6 221 4 US-09-489-039A-7778 Sequence 7778, Ap

ALIGNMENTS

RESULT 1

US-09-198-452A-616
; Sequence 616, Application US/09198452A
; Patent No. 6559294

; GENERAL INFORMATION:

; APPLICANT: Griffais, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

; FILE REFERENCE: 9710-003-999 and treatment of infection

; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 616

; LENGTH: 195

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-198-452A-616

Query Match 44.5%; Score 49; DB 4; Length 195;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22

Db 37 RQKTFLLSHRVLARPSFPVD 56

RESULT 2

US-09-543-681A-5383
; Sequence 5383, Application US/09543681A
; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 5383

; LENGTH: 255

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-5383

Query Match 39.1%; Score 43; DB 4; Length 255;
Best Local Similarity 43.5%; Pred. No. 20;

XX 09-OCT-2001 (first entry)
 DT Peptide #1515 encoded by probe for measuring breast gene expression.
 XX
 DE Probe; human; breast disease; breast cancer; development disorder;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200157270-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 29-JAN-2001; 2001WO-US000661.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-476286/51.
 XX
 PT Novel single exon nucleic acid probe used to measuring gene expression in
 PT a human breast.
 XX
 PS Claim 27; SEQ ID NO 11573; 322pp; English.
 XX
 CC The present invention relates to novel single exon nucleic acid probes
 CC (see AA100010-AA110067). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for measuring human gene expression in
 CC a human breast sample, where the probe hybridises at high stringency to a
 CC nucleic acid expressed in the human breast. The probes are useful for
 CC predicting, diagnosing, grading, staging, monitoring and prognosing
 CC diseases of the human breast, particularly those diseases with polygenic
 CC aetiology. The diseases include: breast cancer, disorders of development,
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative
 CC breast disease and non-carcinoma tumours. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 50 AA;
 SQ
 Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RKKTFLSVHEVMAVENYAKD 22
 |||||
 Db 1 RKKTFLSVHEVMAVENYAKD 20

Search completed: March 18, 2004, 14:14:57
 Job time : 20.9484 secs

CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention

XX Sequence 50 AA;

Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
 |||||
 Db 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 13
 AAM54871
 ID AAM54871 standard; protein; 50 AA.

XX AC AAM54871;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26976.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.

XX Example 4; SEQ ID NO 26976; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention

XX Sequence 50 AA;

Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
 |||||
 Db 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 14

ABG48915
 ID ABG48915 standard; peptide; 50 AA.

XX AC ABG48915;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 27563.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000664.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.

XX Claim 27; SEQ ID NO 27563; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification for complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 50 AA;

Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHEVMAVENYAKD 22
 |||||
 Db 1 RKKTFLSVHEVMAVENYAKD 20

RESULT 15

AAM02833
 ID AAM02833 standard; protein; 50 AA.

XX AC AAM02833;

```
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
XX Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and B7 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 50 AA;

Query Match          91.8%; Score 101; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFSLVHEVMAVENYAKD 22
DB 1 RKKTFSLVHEVMAVENYAKD 20

RESULT 11
ABB19524
ID ABB19524 standard; protein; 50 AA.
AC ABB19524;
DT
DT 23-JAN-2002 (first entry)
DE
DE Protein #1523 encoded by probe for measuring heart cell gene expression.
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
OS Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
```

```
DR WPI; 2001-488999/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
XX Claim 15; SEQ ID NO 21294; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA411305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 50 AA;

Query Match          91.8%; Score 101; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFSLVHEVMAVENYAKD 22
DB 1 RKKTFSLVHEVMAVENYAKD 20

RESULT 12
AAM67252
ID AAM67252 standard; protein; 50 AA.
XX
XX AAM67252;
XX
XX 06-NOV-2001 (first entry)
DT
DE
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
```

PN WO200157272-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US0000663.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX
 XX Claim 27; SEQ ID NO 27814; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP:
 CC see AAI31315-AI57546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 XX human genetic disorders
 XX Sequence 50 AA;
 SQ
 Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RKKTFLSVHEVMAVENYAKD 22
 Db 1 RKKTFLSVHEVMAVENYAKD 20
 RESULT 9
 ABB32389
 ID ABB32389 standard; peptide; 50 AA.
 XX
 XX ABB32389;
 AC
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Peptide #5040 encoded by breast cell single exon nucleic acid probe.
 XX Human; microarray; single exon probe; gene expression; breast; disease;
 KW cancer.
 KW
 XX Homo sapiens.
 OS
 XX WO200157271-A2.
 PN
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US0000662.
 XX
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-496933/54.
 XX New spatially-addressable set of single exon nucleic acid probes, useful
 PT for measuring gene expression in sample derived from human breast,
 PT comprises number of single exon nucleic acid probes.
 XX
 XX Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a peptide
 CC encoded by a single exon nucleic acid probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 50 AA;
 SQ
 Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RKKTFLSVHEVMAVENYAKD 22
 Db 1 RKKTFLSVHEVMAVENYAKD 20
 RESULT 10
 ABB28913
 ID ABB28913 standard; peptide; 50 AA.
 XX
 XX ABB28913;
 AC
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.
 XX Human; microarray; single exon probe; gene expression; breast; disease;
 KW cancer.
 KW
 XX Homo sapiens.
 OS
 XX WO200157271-A2.
 PN
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US0000662.
 XX
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;

Query Match 91.8%; Score 101; DB 5; Length 36;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAKD 22
 DB 1 RKKTFLSVHVMVAVENYAKD 20

RESULT 6
 AAM15096
 ID AAM15096 standard; protein; 50 AA.
 XX AC AAM15096;
 XX DT 12-OCT-2001 (first entry)
 XX DE Peptide #1530 encoded by probe for measuring cervical gene expression.
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
 XX KW cervical cancer.
 XX OS Homo sapiens.
 XX PN WO200157278-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US0000670.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX PF WPI; 2001-488901/53.
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX PT gene expression in human cervical epithelial cells.
 XX PS Claim 27; SEQ ID NO 19922; 487pp; English.
 XX CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs; see AAI10068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 50 AA;
 Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAKD 22
 DB 1 RKKTFLSVHVMVAVENYAKD 20

RESULT 7

ABB34086
 ID ABB34086 standard; peptide; 50 AA.
 XX AC ABB34086;
 XX DT 04-FEB-2002 (first entry)
 XX DE Peptide #1592 encoded by human foetal liver single exon probe.
 XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX OS Homo sapiens.
 XX PN WO200157277-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US0000669.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX PF WPI; 2001-483447/52.
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX PT gene expression in human fetal liver.
 XX PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.
 XX CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 50 AA;
 Query Match 91.8%; Score 101; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKTFLSVHVMVAVENYAKD 22
 DB 1 RKKTFLSVHVMVAVENYAKD 20

RESULT 8
 AAM27545
 ID AAM27545 standard; protein; 50 AA.
 XX AC AAM27545;
 XX DT 17-OCT-2001 (first entry)
 XX DE Peptide #1592 encoded by probe for measuring placental gene expression.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX KW genetic disorder.
 XX OS Homo sapiens.
 XX

CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 117 AA;
 Query Match 100.0%; Score 110; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QARKKTFSLVHEVMAVENYAKD 22
 DB 4 QARKKTFSLVHEVMAVENYAKD 25
 RESULT 4
 AAU79852
 ID AAU79852 standard; protein; 137 AA.
 XX
 AC AAU79852;
 DT 15-JUL-2002 (first entry)
 DE Human cystatin-8 (Zcys8).
 DE Human cystatin-8 (Zcys8) antigenic fragment #4.
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation.
 XX Homo sapiens.
 OS
 PN WO200220567-A2.
 PD 14-MAR-2002.
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Holloway JL, Gao Z, Bishop PD;
 PI WPI; 2002-383044/41.
 DR N-PSDB; ABK49522.
 XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 93-94; 100pp; English.
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This is the amino acid sequence of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 137 AA;
 Query Match 100.0%; Score 110; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.7e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QARKKTFSLVHEVMAVENYAKD 22
 DB 24 QARKKTFSLVHEVMAVENYAKD 45
 RESULT 5
 AAU79856
 ID AAU79856 standard; peptide; 36 AA.
 XX
 AC AAU79856;
 DT 15-JUL-2002 (first entry)
 DE Human cystatin-8 (Zcys8) antigenic fragment #4.
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX Homo sapiens.
 OS
 PN WO200220567-A2.
 PD 14-MAR-2002.
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Holloway JL, Gao Z, Bishop PD;
 PI WPI; 2002-383044/41.
 DR Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 95; 100pp; English.
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 XX
 SQ Sequence 36 AA;

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful for detecting whether a subject's
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 22 AA;
 Query Match 100.0%; Score 110; DB 5; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.3e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QARKKTFSLVHEVMAVENYAKD 22
 Db 1 QARKKTFSLVHEVMAVENYAKD 22
 RESULT 2
 AAU79853
 ID AAU79853 standard; protein; 115 AA.
 XX
 AC AAU79853;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #1.
 XX
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026968.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 94; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC expression of a Zcys8 gene in a biological sample and Zcys8

CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 115 AA;
 Query Match 100.0%; Score 110; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QARKKTFSLVHEVMAVENYAKD 22
 Db 2 QARKKTFSLVHEVMAVENYAKD 23
 RESULT 3
 AAU79854
 ID AAU79854 standard; protein; 117 AA.
 XX
 AC AAU79854;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.
 XX
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026968.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 94-95; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 19,9484 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-5

Perfect score: 110

Sequence: 1 QARKKTFSLVHEVMAVENYAKD 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	110	100.0	22	5 AAU79855	Aau79855 Human cys
2	110	100.0	115	5 AAU79853	Aau79853 Human cys
3	110	100.0	117	5 AAU79854	Aau79854 Human cys
4	110	100.0	137	5 AAU79852	Aau79852 Human cys
5	101	91.8	36	5 AAU79856	Aau79856 Human cys
6	101	91.8	50	4 AAM15096	Aam15096 Peptide #
7	101	91.8	50	4 ABB34086	Abb34086 Peptide #
8	101	91.8	50	4 AAM27545	Aam27545 Peptide #
9	101	91.8	50	4 ABB32389	Abb32389 Peptide #
10	101	91.8	50	4 ABB28913	Abb28913 Peptide #
11	101	91.8	50	4 ABB19524	Abb19524 Protein #
12	101	91.8	50	4 AAM67252	Aam67252 Human bon
13	101	91.8	50	4 AAM54871	Aam54871 Human bra
14	101	91.8	50	4 ABB48915	Abb48915 Human liv
15	101	91.8	50	4 AAM02833	Aam02833 Peptide #
16	101	91.8	50	5 ABB36903	Abb36903 Human pep
17	53	48.2	184	5 ABB30713	Abb30713 Chlamydia
18	49	44.5	195	2 AAY35198	Aay35198 Chlamydia
19	45	40.9	517	4 ABB57957	Abb57957 Drosophil
20	44	40.0	180	6 ADB95534	Adb95534 Allostoc
21	44	40.0	296	4 AAU07104	Aau07104 Human nov
22	44	40.0	318	4 AAU07106	Aau07106 Human nov
23	44	40.0	370	5 ABB97599	Abb97599 Novel hum
24	44	40.0	375	4 AAU07109	Aau07109 Human nov
25	44	40.0	473	4 AAU07111	Aau07111 Human nov

ALIGNMENTS

RESULT 1

AAU79855
ID AAU79855 standard; peptide; 22 AA.

XX AAU79855;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #3.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic fragment.

XX Homo sapiens.

XX WO200220567-A2.

PD 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026868.

PR 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 95; 100pp; English.

CC The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (i) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(i) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large

Adc51944 Human pos
Aau07115 Human nov
Ada55525 Human pro
Adc10176 Human NOV
Aau07119 Human nov
Aau07123 Human kin
Aae16270 Human mem
Abb79941 Human pos
Adc51943 Human pos
Aau20555 Human sec
Aaw98554 H. pylori
Adu37832 Protein e
Ade52552 Caenorhab
Aau33499 Enterococ
Abu29638 Protein e
Aau35034 Enterococ
Abu4481 Protein e
Adc96593 E. faeciu
Abu33247 Protein e
Abb17531 Human ner

26 44 40.0 491 7 ADC51944
27 44 40.0 520 4 AAU07115
28 44 40.0 542 6 ADA55525
29 44 40.0 549 7 ADC10176
30 44 40.0 560 4 AAU07119
31 44 40.0 576 4 AAU07123
32 44 40.0 576 5 AAE16270
33 44 40.0 576 5 ABB79941
34 44 40.0 576 7 ADC51943
35 44 40.0 580 4 AAU20555
36 43 39.1 259 2 AAW98554
37 43 39.1 431 6 ADU37832
38 43 39.1 467 7 ADE52552
39 43 39.1 1054 4 AAU33499
40 43 39.1 1175 6 ABU29638
41 43 39.1 1208 4 AAU35034
42 43 39.1 1217 6 ABU4481
43 43 39.1 1224 7 ADC96593
44 42.5 38.6 281 6 ABU33247
45 42 38.2 139 4 ABB17531

Search completed: March 18, 2004, 14:21:23
Job time : 70.1071 secs

```
RA Ikegawa S., Nakamura Y.;
RT "DD72, a novel mouse gene implicated in the early stage of ectopic
RT ossification.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036743; BAA95411.1; -.
DR EMBL; BC048364; AAH48364.1; -.
DR HSSP; P01034; IG96.
DR MGD; MGI:1930004; Cst10.
DR GO; GO:0004869; F:Cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR001713; Stefina.
DR Pfam; PF00031; cystatin; 1.
DR PRINTS; PR00295; STEFINA.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
SQ SEQUENCE 148 AA; 16451 MW; 637534CBFCSAA179 CRC64;

Query Match 17.7%; Score 113; DB 11; Length 148;
Best Local Similarity 28.2%; Pred. No. 0.00031;
Matches 29; Conservative 18; Mismatches 46; Indels 10; Gaps 3;

QY 15 EVMAVENYAKDSLQWITDQYNKESDDKYHPRIFRVLKVRQVTDHLEYHLNVEMQWTTCC 74
DB 50 EVQKVKFA-----VRTYNDMDNDLYLSKPIRLMSASQVVAGKNYYLKIELGRTTCT 102

QY 75 KPET--TNCVPQER-ELHKQVNCFFSVFVFPWFQYKILNKSC 114
DB 103 KTESNLVDCPFNEQDQKRVICNFQINVPWLNKMSMTNFNC 145

RESULT 15
Q8K397 PRELIMINARY; PRT; 146 AA.
AC Q8K397;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RIKEN cDNA 1110017E11 gene (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027680; AAH27680.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
FT NON TER 1
SQ SEQUENCE 146 AA; 16380 MW; 9D77BB9A6063A5C4 CRC64;

Query Match 17.6%; Score 112.5; DB 11; Length 146;
Best Local Similarity 26.1%; Pred. No. 0.00034;
Matches 23; Conservative 20; Mismatches 38; Indels 7; Gaps 1;

QY 34 YNKESDDKYHPRIFRVLKVRQVTDHLEYHLNVEMQWTTCKP-----ETTNCVPQER 86
DB 56 YNMGSDSLYYFRDTKVIDAKYQLVAGIKYYLTDIESTECKRTRVSGEHMDLTTCPLAAG 115

QY 87 ELHKQVNCFFSVFVFPWFQYKILNKSC 114
DB 116 GQEKLRNCFNLELVFPWNTTQLLKDC 143
```

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK020193; BAB32024.1; -

DR HSSP; P01034; IC96.

DR MGD; MGI:1925859; 8030411P24Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

DR SEQUENCE 130 AA; 14947 MW; DD2F930864B4E584 CRC64;

Query Match 18.9%; Score 121; DB 11; Length 130;

Best Local Similarity 27.5%; Pred. No. 3.8e-05;

Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 8 KTFSLVHEVMAVENYAKDSQWITDQYNKESDDKYHFRIFRLVKVQVTDHLYHLNVE 67

Db 23 KEFLDVTKDL---DYFVASEFAVAQFNNDNNPEENTYKLLLEVGRQKK-TWTMIFLMDLE 78

QY 68 MQWTTCKPBT--NCVPQRELHKQVNCFFSVFVAPWPFQYKILNKSC 114

Db 79 MGRITCKKHENHNCPLQOSREKKVHCVFQVDARPFWSHFILTSTC 127

RESULT 12

Q8VII3 PRELIMINARY; PRT; 130 AA.

AC Q8VII3

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Cystatin SC.

GN 8030411P24Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF440735; AAL30841.1; -

DR MGD; MGI:1925859; 8030411P24Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

DR SEQUENCE 130 AA; 15076 MW; DD34930864AFES9F CRC64;

Query Match 18.9%; Score 121; DB 11; Length 130;

Best Local Similarity 27.5%; Pred. No. 3.8e-05;

Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 8 KTFSLVHEVMAVENYAKDSQWITDQYNKESDDKYHFRIFRLVKVQVTDHLYHLNVE 67

Db 23 KEFLDVTKDL---DYFVASEFAVAQFNNDNNPEENTYKLLLEVGRQKK-TWTMIFLMDLE 78

QY 68 MQWTTCKPBT--NCVPQRELHKQVNCFFSVFVAPWPFQYKILNKSC 114

Db 79 MGRITCKKHENHNCPLQOSREKKVHCVFQVDARPFWSHFILTSTC 127

RESULT 13

Q9DAN8 PRELIMINARY; PRT; 128 AA.

AC Q9DAN8

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE 1700006F03Rik protein (Cystatin TE-1).

GN 1700006F03RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadofa K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bargh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK005670; BAB24179.1; -

DR EMBL; AF440737; AAL30843.1; -

DR MGD; MGI:1916612; 1700006F03Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; Cy; 1.

DR SEQUENCE 128 AA; 15036 MW; 40AFD0010355ED6 CRC64;

Query Match 17.9%; Score 114.5; DB 11; Length 128;

Best Local Similarity 29.7%; Pred. No. 0.00018;

Matches 27; Conservative 20; Mismatches 41; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKESDDKYHFRIFRLVKVQVTDHLYHLNVEQWTTCK--PETTNCVP 83

Db 36 SVEHVVFHFNENQDDDFAYKFLVRRLRQKYT-LKYLVDLEMGRTLCGKYDEIDNCPL 94

QY 84 QRELHKQVNCFFSVFVAPWPFQYKILNKSC 114

Db 95 QEGPGRKVRCTIVTEAWTKFTILNSTC 125

RESULT 14

Q9JMB4 PRELIMINARY; PRT; 148 AA.

AC Q9JMB4

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE DD72 protein (similar to cystatin 10) (Chondrocytes).

GN CST10 OR DD72.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  Bai J., Lao H., Ye X., Li Y., Lou J.;
RT  "Molecular cloning and sequence analysis of cystatin cDNA from two
RL  species of sturgeons.";
DR  EMBL; AF334610; AAK16731.1; -.
DR  HSSP; P01038; 1A90.
DR  GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR  InterPro; IPR000010; Cystatin.
DR  Pfam; PF00031; cystatin; 1.
DR  SMART; SM00043; CY; 1.
DR  PROSITE; PS00287; CYSTATIN; 1.
FT  NON TER
SQ  SEQUENCE 112 AA; 12231 MW; 48CEBPEBDA08C00 CRC64;

Query Match 21.7%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 3.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 20 ENYKDSLOWITDQYNKESDDKYHFRIFRVLKQVQRTDHLHYLVNEMQWTTTCQK---P 76
Db 13 EEGVQDALKFAVAEFNKAQNDMTIHRVSKVKQVQVAGIKYIVTQMRTSCRGGAE 72

QY 77 ETTNC-----VPOBELHKLQVNCFFSVFVAPWPFQYKILNKSCS 115
Db 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNKTCT 112

RESULT 9
Q98SR3
ID Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser schrenckii (Amur sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
DR EMBL; AF334611; AAK16732.1; -.
DR HSSP; P01038; 1A90.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBPEBDA08C00 CRC64;

Query Match 21.7%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 3.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 20 ENYKDSLOWITDQYNKESDDKYHFRIFRVLKQVQRTDHLHYLVNEMQWTTTCQK---P 76
Db 13 EEGVQDALKFAVAEFNKAQNDMTIHRVSKVKQVQVAGIKYIVTQMRTSCRGGAE 72

QY 77 ETTNC-----VPOBELHKLQVNCFFSVFVAPWPFQYKILNKSCS 115
Db 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNKTCT 112

RESULT 10
Q98SR3
ID Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser schrenckii (Amur sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
DR EMBL; AF334611; AAK16732.1; -.
DR HSSP; P01038; 1A90.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBPEBDA08C00 CRC64;

Query Match 21.7%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 3.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY 20 ENYKDSLOWITDQYNKESDDKYHFRIFRVLKQVQRTDHLHYLVNEMQWTTTCQK---P 76
Db 13 EEGVQDALKFAVAEFNKAQNDMTIHRVSKVKQVQVAGIKYIVTQMRTSCRGGAE 72

QY 77 ETTNC-----VPOBELHKLQVNCFFSVFVAPWPFQYKILNKSCS 115
Db 73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLKVNKTCT 112

RESULT 10
```

```

Q8VIH8
ID Q8VIH8 PRELIMINARY; PRT; 130 AA.
AC Q8VIH8;
DT 01-WAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin SC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RA Li Y., Friel P.J., Griswold M.D.;
RT "Molecular cloning and characterization of cystatin SC and cystatin
RL TE-1, new members of the cystatin family.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF442405; AAL35350.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 130 AA; 14981 MW; 7A752359860989C9 CRC64;

Query Match 19.1%; Score 122; DB 11; Length 130;
Best Local Similarity 27.5%; Pred. No. 3e-05;
Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 8 KTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQRTDHLHYLVNVE 67
Db 23 KEFLDVTKDL---DYFVASVFQVAFQNDNNSEENTYRLLEVGRAQKK-TWTMFLMDLE 78

QY 68 MOWTTTCQKPETT--NCVPOBELHKLQVNCFFSVFVAPWPFQYKILNKSC 114
Db 79 MORTICKKHDEINHNCPLLOGSGEKKVCHVQVQDAPWFSHFIVLTSTC 127

RESULT 11
Q9CX46
ID Q9CX46 PRELIMINARY; PRT; 130 AA.
AC Q9CX46;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 8030411P24RIK protein.
GN 8030411P24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bares G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garioldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohteki S.,
RA Hayaishizaki Y.,
```

AC	Q98Y72;
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Cystatin-like 1.
DE	Mus musculus (Mouse).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Testicle;
RC	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA	Datchenko L., Marudian K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahney J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.B.,
RA	Jones S.J., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RN	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Testicle;
RA	Strausberg R.;
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL: BC048646; AAH48646.1; -.
DR	GO: GO:0004869; F.cysteine protease inhibitor activity; IEA.
DR	InterPro: IPR000010; Cystatin.
DR	InferPro: IPR003243; Cystatin_C/M.
DR	Pfam: PF00031; cystatin; 1.
DR	ProDom: PD001231; Cystatin_C/M; 1.
DR	SMART: SM00043; CY; 1.
SQ	SEQUENCE 140 AA; 16199 MW; 32633E99C4697DA0 CRC64;
	Query Match 23.4%; Score 149.5; DB 11; Length 140;
	Best Local Similarity 30.4%; Pred. No. 3.8e-08;
	Matches 31; Conservative 27; Mismatches 41; Indels 3; Gaps 2;
Qy	15 EVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQGVTDHLEYHLNVEMQTTCQ 74
Db	37 KAMSKKN-INSTLFFIRSYNNASNDTYLVQVKLIQQGQLTGTGYLYVTVIGRTPCK 95
Qy	75 KPET--TNCVPQRRLHKONCFVSFAVPWFQYKILNKSC 114
Db	96 KNITKASCPLQSKSLCKSLIYSVPMWNYQLWNNSC 137
RESULT 8	
Q98SR4	ID Q98SR4 PRELIMINARY; PRT; 112 AA.
AC	Q98SR4;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Cystatin (Fragment).
OS	Acipenser sinensis (Chinese sturgeon).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC	Acipenser.
OX	NCBI TaxID=61970;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 70.1071 Seconds
(without alignments)
526.560 Million cell updates/sec

Title: US-09-941-314-4
Perfect score: 640
Sequence: 1 LPQARKKTFLSVHEVWAVE.....VFAPVFEQYKILNKSCSSD 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvms:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	640	100.0	138	4 Q8WXU6	Q8wxu6 homo sapien
2	411.5	64.3	103	4 Q8WXU5	Q8wxu5 homo sapien
3	360	56.2	139	11 Q8K5A3	Q8k5a3 rattus norv
4	188.5	29.5	141	11 Q9DAP1	Q9dap1 mus musculu
5	188.5	29.5	141	11 Q80ZN5	Q80zn5 mus musculu
6	166.5	26.0	140	11 Q9EPX9	Q9epx9 mus musculu
7	149.5	23.4	140	11 Q80Y72	Q80y72 mus musculu
8	139	21.7	112	13 Q98SR4	Q98sr4 acipenser s
9	139	21.7	112	13 Q98SR3	Q98sr3 acipenser s
10	122	19.1	130	11 Q8VIH8	Q8vih8 rattus norv
11	121	18.9	130	11 Q9CX46	Q9cx46 mus musculu
12	121	18.9	130	11 Q8VII3	Q8vii3 mus musculu
13	114.5	17.9	128	11 Q9DAN8	Q9dan8 mus musculu
14	113	17.7	148	11 Q9JMH4	Q9jmh4 mus musculu
15	112.5	17.6	146	11 Q8K397	Q8k397 mus musculu
16	112.5	17.6	149	11 Q8VHC1	Q8vhc1 rattus norv

17	112.5	17.6	149	11 Q9D1B1	Q9d1b1 mus musculu
18	109	17.0	167	4 Q7Z4J8	Q7z4j8 homo sapien
19	105	16.4	128	11 Q8VII2	Q8vii2 rattus norv
20	98.5	15.4	144	13 Q8JFUS	Q8jfu5 brachydanio
21	98	15.3	167	11 Q9QWLS	Q9qw15 mus musculu
22	95.5	14.9	109	5 Q9TY65	Q9ty65 onchocerca
23	92.5	14.5	161	5 Q16159	Q16159 brugia mala
24	90.5	14.1	148	5 Q9NH95	Q9nh95 litomosoides
25	86	13.4	133	11 Q9D264	Q9d264 mus musculu
26	84.5	13.2	125	5 Q25620	Q25620 onchocerca
27	84	13.1	127	5 P90698	P90698 brugia mala
28	84	13.1	425	3 Q12700	Q12700 debaryomyce
29	82.5	12.9	157	5 Q17108	Q17108 acanthochei
30	82	12.8	127	5 Q9U9A1	Q9u9a1 onchocerca
31	80.5	12.6	133	5 Q8MVB6	Q8mvb6 ixodes scap
32	80	12.5	498	5 Q16454	Q16454 caenorhabdi
33	79.5	12.4	2773	5 Q8IBF4	Q8ibf4 plasmodium
34	79	12.3	430	11 Q63581	Q63581 rattus norv
35	78	12.2	996	4 Q8NDM7	Q8ndm7 homo sapien
36	77.5	12.1	115	6 Q95K43	Q95k43 macaca fasc
37	76.5	12.0	462	13 Q7ZY91	Q7zy91 xenopus lae
38	75.5	11.8	209	5 Q815X0	Q815x0 plasmodium
39	75.5	11.8	465	13 Q801E5	Q801e5 xenopus lae
40	75	11.7	140	6 Q7YRP6	Q7yrp6 sus scrofa
41	75	11.7	400	4 Q8WXX6	Q8wxk6 homo sapien
42	75	11.7	400	4 Q8WXX7	Q8wxk7 homo sapien
43	75	11.7	423	11 P70517	P70517 rattus norv
44	75	11.7	891	5 Q8IB61	Q8ib61 plasmodium
45	74.5	11.6	658	16 Q81160	Q81160 bacillus ce

ALIGNMENTS

RESULT 1

Q8WXU6 PRELIMINARY; PRT; 138 AA.

AC Q8WXU6
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SC13.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;

RT "SC13: A novel epididymal specific member of the cystatin family.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF335480; AAL71991.1; -

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR00010; Cystatin.

DR Pfam; PF00031; Cystatin; 1.

DR SMART; SM00043; CY; 1.

SQ SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;

Query Match 100.0%; Score 640; DB 4; Length 138;

Best Local Similarity 100.0%; Pred. No. 2.9e-60;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPQARKKTFLSVHEVWAVENYAKDSLOWITDQYNKESDDKYHFRFRVLKQVRQVTDHL 60

DB 22 LPYQARKKTFLSVHEVWAVENYAKDSLOWITDQYNKESDDKYHFRFRVLKQVRQVTDHL 81

QY 61 EYHLNVEMQWTTCKPETTNCVQRELHKQVNCFFSVFAPVFEQYKILNKSCSSD 117

DB 82 EYHLNVEMQWTTCKPETTNCVQRELHKQVNCFFSVFAPVFEQYKILNKSCSSD 138

RESULT 2

Q8WXU5

Search completed: March 18, 2004, 14:16:08
Job time : 13.9286 secs

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehtra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas J.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [7]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klugner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie H.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE OF 27-146.
RX MEDLINE=82222268; PubMed=6283552;
RA Grubb A., Loeffberg H.;
RT "Human gamma-trace, a basic microprotein: amino acid sequence and
presence in the adenohypophysis.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).
RN [9]
RP SEQUENCE OF 27-73.
RX MEDLINE=84110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
RA Machleidt W.;
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
of cystatin from chicken egg white.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [10]
RP SEQUENCE OF 27-76.
RX MEDLINE=84128015; PubMed=6365094;
RA Brzin J., Popovic T., Turk V.;
RT "Human cystatin, a new protein inhibitor of cysteine proteinases.";
RL Biochem. Biophys. Res. Commun. 118:103-109(1984).
RN [11]
RP DISULFIDE BONDS.
RA Grubb A., Loeffberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
cystatin.";
RL FEBS Lett. 170:370-374(1984).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.
RX MEDLINE=21173909; PubMed=11276250;
RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,
RA Abrahamson M., Jaskolski M.;
RT "Human cystatin C, an amyloidogenic protein, dimerizes through
three-dimensional domain swapping.";
RL Nat. Struct. Biol. 8:316-320(2001).
RN [13]
RP VARIANT GLN-94.
RX MEDLINE=92316504; PubMed=1352269;
RA Abrahamson M., Jonsdottir S., Olafsson I., Jenson O., Grubb A.;
RT "Hereditary cystatin C amyloid angiopathy: identification of the
disease-causing mutation and specific diagnosis by polymerase chain
reaction based analysis.";
RL Hum. Genet. 89:377-380(1992).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
thought to serve an important physiological role as a local
regulator of this enzyme activity.
CC -!- SUBUNIT: Homodimer.
CC -!- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,
vas deferens, brain, thymus, and ovary and the lowest in the
submandibular gland.
CC -!- DISEASE: Defects in CST3 are a cause of hereditary cerebral
hemorrhage with amyloidosis (HCHWA) [MIM:105150]; also known as
cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis
Icelandic type. HCHWA is characterized by a thickening of the
cerebral arteries walls with deposition of material with the
characteristics of amyloid.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05607; CAA29096.1; -;
DR EMBL; X52255; CAA36497.1; -;
DR EMBL; M27891; AAA52164.1; -;
DR EMBL; M27889; AAA52164.1; JOINED.
DR EMBL; M27890; AAA52164.1; JOINED.
DR EMBL; X61681; CAA43856.2; -;
DR EMBL; X61682; CAA43856.2; JOINED.
DR EMBL; X61683; CAA43856.2; JOINED.
DR EMBL; AF319564; AAK11570.1; -;
DR EMBL; AL121894; CAC05424.1; -;
DR EMBL; BC013083; AAH13083.1; -;
DR PIR; S10216; UDHU.
DR PDB; 1G96; 06-APR-01.
DR Genew; HGNC:2475; CST3.
DR MIM; 604312; -;
DR MIM; 105150; -;
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109
FT DISULFID 123 143
Query Match 24.2%; Score 155; DB 1; Length 146;
Best Local Similarity 33.7%; Pred. No. 5.6e-09;
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;
QY 20 ENYAKDSLOWITDQYNKESDDKYHFRIVLVKQVQVTDHLYHVLNVMQWTTCK--PE 77
DB 46 EEGVRALDFAVGEYKASNDYHSALQVVRARKQIVAGVNYFLDELGRITCTKTPN 105
QY 78 TTNC-VPQERELHKQNCFFSFAVFWFQYKILNKSQSSD 117
DB 106 LDCPPHDPHLKRAFCFSQIYAVPW-OQTMTLSKSTCOD 145
RESULT 15
CYTC_MACMU
ID _CYTC_MACMU STANDARD; PRT; 146 AA.

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leivaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McMay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Suze C.D., Smith M.L., Soderlund C., Steward C.A., Sultston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -I- SUBCELLULAR LOCATION: Secreted (Potential).
CC -I- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL096677; CAC03440.2; -
DR HSSP: P01038; J490
DR Genew; HGNC:15958; CST11.
DR InterPro; IPR000010; Cystatin.
DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 165 CYSTATIN-LIKE 1.
FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 111 121 BY SIMILARITY.
FT DISULFID 134 154 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 165 AA; 19312 MW; 9D66D685875DAEEA CRC64;
Query Match 24.6%; Score 157.5; DB 1; Length 165;
Best Local Similarity 26.2%; Pred. No. 3.6e-09;
Matches 32; Conservative 30; Mismatches 37; Indels 23; Gaps 3;
QY 15 EVMAVENYAKDSQWITDQYKNSDDKYHFRIFRLVKVQ-----R 54
Db 34 KLSKSN-MNSTLNFFIQSTNNASNDYLRVRLRSQMRVSHWMLGVHNTSTDSR 92
QY 55 QVTDHLEYHLNMQWTTQKPTTN--CVPQRELHKQVNCFFSFAVPWFQYKILNK 112
Db 93 QLITGVYEIVTVKIGTWKCKRNDTSNCCPLQSKKLRLKSLICSLIIVTWPWVFLWN 152
QY 113 SC 114
Db 153 SC 154

RESULT 14
CYTC_HUMAN STANDARD; PRT; 146 AA.
AC P01034;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
DE (Post-gamma-globulin).
GN CST3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=87219149; PubMed=3495457;
RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;
RT "Molecular cloning and sequence analysis of cDNA coding for the
RT precursor of the human cysteine proteinase inhibitor cystatin C.";
RL FEBS Lett. 216:229-233(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=90303202; PubMed=2363674;
RA Abrahamson M., Olafsson I., Paledottir A., Ulvsbaeck M., Lundwall A.,
RA Jansson O., Grubb A.;
RT "Structure and expression of the human cystatin C gene.";
RL Biochem. J. 268:287-294(1990).
RN [3]
RP SEQUENCE FROM N.A. (HCHWA VARIANT).
RC TISSUE=Brain;
RX MEDLINE=89235594; PubMed=2541223;
RA Levy E., Lopez-Otin C., Ghiso J., Gelnert D., Frangione B.;
RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is
RT related to a mutation in the cystatin C gene, an inhibitor of
RT cysteine proteases";
RL J. Exp. Med. 169:1771-1778(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89350949; PubMed=2764935;
RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,
RA Isemura S., Sanada K.;
RT "The human cystatin C gene (CST3) is a member of the cystatin gene
RT family which is localized on chromosome 20.";
RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
RN [5]
RP SEQUENCE FROM N.A.
RX Dickinson D.P., Hewett-Emmett D., Thiesse M.;
RT "Acquisition of complex patterns of differential expression in
RT epithelial cell populations during the evolution of type 2 cystatin
RT genes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leivaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McMay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Buffalo;
 RX MEDLINE=90092122; PubMed=2689174;
 RA Cole T., Dickson P.W., Ennard F., Averill F., Risbridger G.,
 RA Gauthier F., Schreiber G.;
 RT "The cDNA structure and expression analysis of the genes for the
 RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
 RT in rat brain."; Eur. J. Biochem. 186:35-42(1989).
 RL [2]
 RN SEQUENCE OF 8-127.
 RP MEDLINE=90380276; PubMed=2400577;
 RX Ennard F., Ennard A., Faucher D., Capony J.-P., Derancourt J.,
 RA Brillard M., Gauthier F.;
 RT "Rat cystatin C: the complete amino acid sequence reveals a site for
 RT N-glycosylation."; J. Biol. Chem. Hoppe-Seyler 371:161-166(1990).
 RL [3]
 RN SEQUENCE OF 8-49.
 RP MEDLINE=89313020; PubMed=3044831;
 RX Ennard A., Ennard F., Faucher D., Gauthier F.;
 RA "Two rat homologues of human cystatin C."; FEBS Lett. 236:475-478(1988).
 RL [4]
 RN SEQUENCE OF 8-20.
 RP TISSUE=Sertoli cells;
 RC MEDLINE=92225121; PubMed=1563513;
 RX Ennard A., Ennard F., Guillou F., Gauthier F.;
 RA "Production of the cysteine proteinase inhibitor cystatin C by rat
 RT Sertoli cells."; FEBS Lett. 300:131-135(1992).
 RL [5]
 CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity. Known to inhibit cathepsin B,
 CC H, and L.
 CC -!- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X16957; CAA34831.1; -
 DR PIR; S07085; S07085.
 DR PIR; S10587; S10587.
 DR HSSP; P01034; I696.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 KW Thiol protease inhibitor; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 1
 FT CHAIN 8 127 CYSTATIN C.
 FT ACT_SITE 18 18 REACTIVE SITE.
 FT SITE 62 66 SECONDARY AREA OF CONTACT.
 FT DISULFID 80 90 BY SIMILARITY.
 FT DISULFID 104 124 BY SIMILARITY.
 FT CONFLICT 25 25 A -> E (IN REF. 2).
 SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925953 CRC64;
 Query Match 26.8%; Score 171.5; DB 1; Length 127;
 Best Local Similarity 33.3%; Pred. No. 9.9e-11;
 Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;
 OX 15 EVMAVNYAKDSIQWITDQYNKESDDKXIFRIFRLVKQRQVTDHLEYHLNVMQWTTQC 74
 DB 22 EADASEGVQRALDFAVSEYKNGSDAYSHRAIQVVRARQLVAGINYILDVEMGRTTCT 81

QY 75 KPET--TNC--VPQRELHKQVNCFFSVFAVPWFQYKILNKSCSS 116
 DB 82 KSQTLNLTNCFFHDQPHLMKALCSQIYSPVWKGTHTLTKSCKN 126
 RESULT 8
 CVT_CHICK
 ID_CVT_CHICK STANDARD; PRT; 139 AA.
 AC P01038;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin precursor (Egg-white cystatin).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90008873; PubMed=2793849;
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
 RT and tissue distribution."; J. Biol. Chem. 264:17164-17169(1989).
 RL [2]
 RN SEQUENCE OF 24-139.
 RP MEDLINE=84178305; PubMed=6712597;
 RX Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
 RA "Cystatin. Amino acid sequence and possible secondary structure."; Biochem. J. 217:813-817(1984).
 RL [3]
 RN SEQUENCE OF 24-139.
 RP MEDLINE=84110059; PubMed=6662498;
 RX Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
 RA Machleidt W.;
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 RT of cystatin from chicken egg white."; Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
 RL [4]
 RN CHARACTERIZATION OF PROTEIN.
 RX MEDLINE=83256421; PubMed=6409085;
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Savers C.A.,
 RA Santer D.C., Barrett A.J.;
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
 RT purification from egg white, characterization, and detection in
 RT chicken serum."; Biochem. J. 211:129-138(1983).
 RL [5]
 RN DISULFIDE BONDS.
 RA Grubb A., Loeffberg H., Barrett A.J.;
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 RT cystatin."; FEBS Lett. 170:370-374(1984).
 RL [6]
 RN PHOSPHORYLATION.
 RX MEDLINE=89252033; PubMed=2721673;
 RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
 RA Bode W.;
 RT "The cysteine proteinase inhibitor chicken cystatin is a
 RT phosphoprotein."; FEBS Lett. 248:162-168(1989).
 RL [7]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=89052676; PubMed=3191914;
 RA Bode W., Eggh R., Musil D., Thiele U., Huber R., Karshikov A.,
 RA Brzin J., Kos J., Turk V.;
 RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
 RT its possible mode of interaction with cysteine proteinases."; EMBO J. 7:2593-2599(1988).
 RL [8]
 RN STRUCTURE BY NMR.
 RP


```

cystatin-related epididymal spermatogenic gene.";
RL Biochem. J. 340:85-93(1999).
CC -1- FUNCTION: Performs a specialized role during sperm development and maturation.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the cystatin family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
EMBL: AF090692; AAC36317.1; -.
DR HSP; P01034; IG96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 19 POTENTIAL
FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC PROTEIN.
FT SITE 77 81 PROTEIN.
FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).
FT FT DISULFID 119 139 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 142 AA; 16246 MW; FB873FAA6B6CAB34 CRC64;
Query Match 31.1%; Score 199; DB 1; Length 142;
Best Local Similarity 35.7%; Pred. No. 1.8e-13;
Matches 41; Conservative 26; Mismatches 36; Indels 12; Gaps 4;
QY 12 SVHEWAVENYAKDSIQ-----WIT-DOYNKESDDKXHFIPVLVKVQRQVTDHLEY 62
DB 25 SKNEVAQRYFGSISISANVKNQCVFPAKKEYNKSEDKYLFLLDKTLHATLQITDRMEY 84
QY 63 HLMVEMQWTTQCRP--ETTNVCVPOER-ELHKVQNCFFSVFAVPWFQYKILNKSC 114
DB 85 HIDVQISRCRKPLNNTENCIPQKNPKLEKLSCLFVGLPFWNEFDLLSKEC 139
-----
RESULT 5
CST8_HUMAN STANDARD; PRT; 142 AA.
AC O60676;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin 8).
DE CST8 OR CRES.
GN Homo sapiens (Human).
OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=95344753; PubMed=7619504;
RA Cornwall G.A., Hann S.R.;
RT "Transient appearance of CRES protein during spermatogenesis and caput epididymal sperm maturation."
RT Mol. Reprod. Dev. 41:37-46(1995).
RL [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA
RX

```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AL096677; CAC13170.1; -
 DR EMBL; AL096677; CAC17423.1; -
 DR HSSP; P01038; 1A90.
 DR Genew; HGNC:15959; CST11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; Cy; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KW Thiol protease inhibitor; Signal; Alternative splicing.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 137 CYSTATIN 11.
 FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 93 101 BY SIMILARITY.
 FT DISULFID 114 134 BY SIMILARITY.
 FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 76 110 Missing (in isoform 2).
 FT /FTID=VSP 001260.
 SQ SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 640; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.6e-58;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRFLVKVQRTVDHL 60
 DB 21 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRFLVKVQRTVDHL 80
 QY 61 EYHLNEMQWTTCKPQETTCVQPERLHKQVNCFFSVFVAPVFPFQYKILNKSCSSD 117
 DB 81 EYHLNEMQWTTCKPQETTCVQPERLHKQVNCFFSVFVAPVFPFQYKILNKSCSSD 137

RESULT 2

CST11_MOUSE STANDARD; PRT; 139 AA.
 AC Q9D269;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin 11 precursor.
 GN CST11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadori K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
 RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).
 CC -I- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -I- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AK020300; BAB32061.1; -
 DR HSSP; P01034; 1G96.
 DR MGD; MGI:1925490; Cst11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; Cy; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KW Thiol protease inhibitor; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 139 CYSTATIN 11.
 FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 94 102 BY SIMILARITY.
 FT DISULFID 115 135 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 139 AA; 16217 MW; P228D9815FA32640 CRC64;

Query Match 58.1%; Score 372; DB 1; Length 139;
 Best Local Similarity 54.9%; Pred. No. 3.7e-31;
 Matches 62; Conservative 29; Mismatches 22; Indels 0; Gaps 0;

QY 3 YQARKKTFSLVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRFLVKVQRTVDHLEY 62
 DB 24 YQVKKTKTFRIBESVSALESSEVKTLEYVTDYENKESDLYNFRILKIMKQVTHLEY 83
 QY 63 HLNEMQWTTCKPQETTCVQPERLHKQVNCFFSVFVAPVFPFQYKILNKSCS 115
 DB 84 HITVEMQRTTCLKTETSLCDIQKGLHKKIQCVFSVYVAPVFPVFKILKQKCT 136

RESULT 3

CST8_MOUSE STANDARD; PRT; 142 AA.
 AC P32766; O89102;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-
 DE related epididymal specific protein) (Cystatin 8).
 GN CST8 OR CR8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H, and CD-1;
 RX MEDLINE=99247899; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.;
 RA "Structure, alternative splicing and chromosomal localization of the
 RA cystatin-related epididymal spermatogenic gene";
 RT Biochem. J. 340:85-93(1999).
 RN [2]
 RP SEQUENCE OF 4-142 FROM N.A.
 RC TISSUE=Epididymis;
 RX MEDLINE=93078799; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;
 RA "The CR8 gene: a unique testis-regulated gene related to the cystatin
 RT family is highly restricted in its expression to the proximal region
 RT of the mouse epididymis";
 RL Mol. Endocrinol. 6:1653-1664(1992).
 CC -I- FUNCTION: Performs a specialized role during sperm development and

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 13.9286 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-4
Perfect score: 640
Sequence: 1 LPYQARKKTFSLVHEVWAVE.....VFAPWFEQYKILNKSCSSD 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	640	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	372	58.1	139	CS11 MOUSE	P01042 mus musculus
3	214	33.4	142	CST8 MOUSE	P08934 rattus norv
4	199	31.1	142	CST8 RAT	O87rl6 vibrrio para
5	188.5	29.5	142	CST8 HUMAN	P08932 rattus norv
6	174.5	27.3	148	CYT8 RABIT	P01048 rattus norv
7	171.5	26.8	127	CYT8 RAT	O08677 mus musculus
8	170.5	26.5	139	CYT CHICK	O74232 kluyveromyc
9	169.5	26.5	116	CYT COTUJA	P44689 haemophilus
10	160.5	25.1	140	CYT8 MOUSE	O7mmw8 vibrrio vuln
11	160.5	25.1	141	CYT8 HUMAN	Q06430 homo sapien
12	160	25.0	146	CYT8 SAISC	P29717 candida alb
13	157.5	24.6	165	CSTL HUMAN	
14	155	24.2	146	CYT8 HUMAN	
15	155	24.2	146	CYT8 MACMU	
16	153.5	24.0	148	CYT8 BOVIN	
17	146.5	22.9	141	CYT8 BOVIN	
18	143	22.3	129	CYT CYPCA	
19	139.5	21.8	141	CYT8 HUMAN	
20	135	21.1	147	CST9 HUMAN	
21	130.5	20.4	142	CYT8 HUMAN	
22	129.5	20.2	141	CYT8 RAT	
23	127	19.8	111	CYT BITAR	
24	127	19.8	130	CYT ONCMY	
25	127	19.8	137	CST9 MOUSE	
26	126	19.7	130	CYT ONCKE	
27	121.5	19.0	149	CYT8 HUMAN	
28	109	17.0	145	CYT8 HUMAN	
29	98.5	15.4	434	KNL2 BOVIN	
30	98.5	15.4	619	KNH2 BOVIN	
31	98	15.3	144	CYT8 BOVIN	
32	96.5	15.1	436	KNL1 BOVIN	
33	96.5	15.1	621	KNH1 BOVIN	

34	95.5	14.9	162	1	CYTX ONCVO	P22085 onchocerca
35	93.5	14.6	644	1	KNG HUMAN	P01042 homo sapien
36	87	13.6	639	1	KNG RAT	P08934 rattus norv
37	82.5	12.9	474	1	SYE_VIBPA	O87rl6 vibrrio para
38	81	12.7	430	1	KNT2 RAT	P08932 rattus norv
39	79	12.3	430	1	KNT1 RAT	P01048 rattus norv
40	79	12.3	661	1	KNG MOUSE	O08677 mus musculus
41	77	12.0	214	1	CAP2 KLULA	O74232 kluyveromyc
42	77	12.0	430	1	MESJ_HAEIN	P44689 haemophilus
43	75.5	11.8	474	1	SYE_VIBVY	O7mmw8 vibrrio vuln
44	75	11.7	400	1	BGIB_HUMAN	Q06430 homo sapien
45	73.5	11.5	438	1	EXG_CANAL	P29717 candida alb

RESULT 1

CS11_HUMAN	STANDARD;	PRT;	137 AA.
AC	Q9H112; Q9H113;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DE	Cystatin 11 precursor.		
GN	CST11 OR CST8L.		
OS	Homo sapiens (Human).		
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RX	MEDLINE=21638749; PubMed=11780052;		
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,		
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,		
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,		
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,		
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,		
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,		
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,		
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,		
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,		
RA	Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,		
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,		
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,		
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,		
RA	Lehvaeslahti M.H., Leveraha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,		
RA	Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,		
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,		
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Beck A.I.,		
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,		
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,		
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,		
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,		
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,		
RA	Whithead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,		
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,		
RA	Rogers J.;		
RT	"The DNA sequence and comparative analysis of human chromosome 20.;"		
RL	Nature 414:865-871(2001).		
CC	-1- SUBCELLULAR LOCATION: Secreted (Potential).		
CC	-1- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=1;		
CC	isoId=Q9H112-1; Sequence=Displayed;		
CC	Name=2;		
CC	isoId=Q9H112-2; Sequence=VSP_001260;		
CC	Note=No experimental confirmation available;		
CC	-1- SIMILARITY: Belongs to the cystatin family.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		

Search completed: March 18, 2004, 14:23:12
Job time : 22.5179 secs

A;Cross-references: GB:J03870; NID:g337751; PIDN:AAA60299.1; PID:g337752
R.;Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988
A;Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily
A;Reference number: S02489; MUID:89076505; PMID:3202964
A;Accession: S02489
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 21-141 <SA2>
R.;Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.
Gene 61, 329-338, 1987
A;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem
A;Reference number: A91589; MUID:88185836; PMID:3446578
A;Accession: A29632
A;Molecule type: DNA
A;Residues: 1-86, 'I', 88-141 <SA1>
R.;Isemura, S.; Saitoh, E.; Sanada, K.
FEBS Lett. 198, 145-149, 1986
A;Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystat
A;Reference number: A01273; MUID:86164938; PMID:3514272
A;Accession: A01273
A;Molecule type: protein
A;Residues: 29-141 <SE>
R.;Ramasubbu, N.; Reddy, M.S.; Bergey, E.J.; Haraszthy, G.G.; Soni, S.D.; Levine, M.J.
Biochem. J. 280, 341-352, 1991
A;Title: Large-scale purification and characterization of the major phosphoproteins and
A;Reference number: S19279; MUID:92082469; PMID:1747107
A;Accession: S19279
A;Status: preliminary
A;Molecule type: protein
A;Residues: 21-55 <RAM>
C;Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a
ences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl
C;Genetics:
A;Gene: GDB:CST1
A;Cross-references: GDB:119815; OMIM:123855
A;Map position: 20p11.2-20p11.2
C;Superfamily: cystatin; cystatin homolog
C;Keywords: cysteine proteinase inhibitor; extracellular protein; saliva
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-141/Product: cystatin SA-I #status experimental <MAT1>
F;29-141/Product: cystatin SN-I #status experimental <MAT2>
F;30-141/Domain: cystatin homolog <CYS>
F;76-80/Region: inhibitory #status predicted
F;94-104,118-138/Disulfide bonds: #status predicted

Query Match 21.8%; Score 139.5; DB 1; Length 141;
Best Local Similarity 29.6%; Pred. No. 8.9e-07;
Matches 29; Conservative 21; Mismatches 45; Indels 3; Gaps 2;
QY 20 ENYAKOSLOWITDOYNKESDDKYHFRIFRLVKVRQVTDHLEYHLNVEQWTTTCQK--PE 77
DB 41 DEWQVRAHFAISYNKATDDYRRLVRLURARQQTGGVNVFFDEVGRTICTSKQPN 100
QY 78 TTNCVPOER-ELHKQVNCFFSVFAPVPPEQYKILNKSC 114
DB 101 LDTCAFHEQPELQKQLCSFEIYEVPPWENRSLVKSRC 138

RESULT 11
JC2040
N;Alternate names: chum salmon
C;Species: Oncorhynchus keta (chum salmon)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 16-Jul-1999
C;Accession: JC2040
R.;Koide, Y.; Noso, T.
BioSci. Biotechnol. Biochem. 58, 164-169, 1994
A;Title: The complete amino acid sequence of pituitary cystatin from chum salmon.
A;Reference number: JC2040; MUID:94162738; PMID:7764512
A;Accession: JC2040
A;Molecule type: protein
A;Residues: 1-111 <KOI>

A;Accession: A01270
A;Molecule type: protein
A;Residues: 27-131,'S',133-146 <GRU>
R;Chieso, J.; Jensen, O.; Frangione, B.
Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986
A;Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland
A;Reference number: A25434; MUID:86206076; PMID:3517880
A;Accession: A25434
A;Molecule type: protein
A;Residues: 37-93,'O',95-146 <GHI>
R;Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983
A;Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin
A;Reference number: S01461; MUID:84110059; PMID:6662498
A;Accession: S12288
A;Molecule type: protein
A;Residues: 27-73 <TUR>
R;Brzin, J.; Popovic, T.; Turk, V.
Biochem. Biophys. Res. Commun. 118, 103-109, 1984
A;Title: Human cystatin, a new protein inhibitor of cysteine proteinases.
A;Reference number: A32732; MUID:84128015; PMID:6365094
A;Accession: A32732
A;Molecule type: protein
A;Residues: 27-76 <BRZ>
R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jensen, O.; Grubb, A.
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst
A;Reference number: A60552; MUID:90193615; PMID:2315647
A;Accession: A60552
A;Molecule type: protein
A;Residues: 27-49,'XX',52-64 <OLA>
A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
e defective gene is not present in CSF but is found instead in amyloid deposits
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
A;Title: Different forms of human cystatin C.
A;Reference number: S10607; MUID:91025625; PMID:2222856
A;Accession: S10607
A;Molecule type: protein
A;Residues: 27-53 <POP>
A;Experimental source: urine, kidney disease
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als
R;Grubb, A.; Lofberg, H.; Barrett, A.J.
FEBS Lett. 170, 370-374, 1984
A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.
A;Reference number: S01462
A;Contents: annotation; disulfide bonds
R;Berti, P.J.; Storer, A.C.
Biochem. J. 302, 411-416, 1994
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility
A;Reference number: S55305; MUID:94379969; PMID:8092991
A;Accession: S55305
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-49;106-146 <BER>
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl
f patients with certain autoimmune diseases.
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importa
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr
C;Genetics:
A;Gene: GDB:CST3
A;Cross-references: GDB:1119817; OMIM:105150
A;Map position: 20p11.2-20p11.2
A;Introns: 81/3, 119/3
C;Superfamily: Cystatin; cystatin homology
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyprolin
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-146/Product: cystatin C #status experimental <MAT>
F;35-146/Domain: cystatin homology <CYS>
F;81-85/Region: inhibitory #status predicted
F;29/Modified site: hydroxyproline (Pro) (partial) #status experimental
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 24.2%; Score 155; DB 1; Length 146;
Best Local Similarity 33.7%; Pred. No. 2.7e-08;
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;
QY 20 ENYAKDSLQWITDQYNKESDDKYHFRIFVLKQVQRTDHLVHLVAVEMQWTTTCQK--PE 77
DB 46 ESGVRALDFAVGEYKNSNDWTHSALQVVRARKQVAGVNYFLDVLGRITCTKQPN 105
QY 78 TTNC-VPQREHLHKQVNCFFSVFAVFWFQYKILNKSCSSD 117
DB 106 LONCPHDQPHLKRKAFCSQIYAVPW-QGTWTLSTKSTCOD 145
RESULT 8
UDBO
cystatin - bovine
N;Alternate names: thiol proteinase inhibitor
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 06-Dec-1996
C;Accession: A01271
R;Hirado, M.; Teunasaawa, S.; Sakiyama, F.; Niinobe, M.; Fujii, S.
FEBS Lett. 186, 41-45, 1985
A;Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase in
A;Reference number: A01271; MUID:85231205; PMID:3891407
A;Accession: A01271
A;Molecule type: protein
A;Residues: 1-112 <HIR>
C;Superfamily: cystatin; cystatin homology
C;Keywords: colostrum; cysteine proteinase inhibitor
F;2-112/Domain: cystatin homology <CYS>
F;48-52/Region: inhibitory #status predicted
F;66-76,90-110/Disulfide bonds: #status predicted
Query Match 24.0%; Score 153.5; DB 1; Length 112;
Best Local Similarity 30.6%; Pred. No. 2.8e-08;
Matches 30; Conservative 23; Mismatches 42; Indels 3; Gaps 2;
QY 20 ENYAKDSLQWITDQYNKESDDKYHFRIFVLKQVQRTDHLVHLVAVEMQWTTTCQKETT 79
DB 13 EEGVQALGFVAVSEFNKRSNDAYQSRVVRVVRARKQVSGMNYFLDVLGRITCTKQAN 72
QY 80 --NC-VPQREHLHKQVNCFFSVFAVFWFQYKILNKSC 114
DB 73 LDSCPPHNPFLKREKLCSCFQVYVWPMNTINLVKFSK 110
RESULT 9
UDHUPI
cystatin S precursor - human
N;Alternate names: cystatin SA-III; salivary acidic protein-1
C;Species: Homo sapiens (man)
C;Date: 25-Feb-1985 #sequence_revision 08-Feb-1996 #text_change 16-Jul-1999
C;Accession: S17667; S16500; A01272; A29603; S19280; A56508
R;Bobek, L.A.; Aguirre, A.; Levine, M.J.
Biochem. J. 278, 627-635, 1991
A;Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ and
A;Reference number: S17667; MUID:91378918; PMID:1898352
A;Accession: S17667
A;Molecule type: mRNA
A;Residues: 1-141 <BOB>
A;Cross-references: EMBL:X54667; NID:G30365; PIDN:CAA38478.1; PID:G30366
R;Lamkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.
Arch. Biochem. Biophys. 288, 664-670, 1991
A;Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle,
A;Reference number: S16500; MUID:91378515; PMID:1898055
A;Accession: S16500
A;Status: preliminary
A;Molecule type: protein
A;Residues: 21-134,'D',136-141 <IHU>
R;Isemura, S.; Saitoh, E.; Sanada, K.
J. Biochem. 96, 489-498, 1984
A;Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole saliv
A;Reference number: A91985; MUID:85054716; PMID:6501254

Db 80 LTNCPPHDPHLMRKALCSFQIYSPVWKGTHTLTKSSCKN 119

RESULT 5

B29632

cystatin SA precursor - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1989 #sequence revision 10-Jun-1989 #text_change 16-Jul-1999

C:Accession: B29632, S02490; A41422; B27015

R:Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A:Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three members

A:Reference number: A91589; MUID:88185836; PMID:3446578

A:Accession: B29632

A:Molecule type: DNA

A:Residues: 1-141 <SAI>

A:Cross-references: GB:M19673; GB:M19170; NID:G186403; PIDN:AAA36116.1; PID:G386826

A>Note: the authors translated the codon GAC for residue 129 as Asn

R:Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A:Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A:Reference number: S02489; MUID:89076505; PMID:3202964

A:Accession: S02490

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 21-141 <SA2>

R:Isemura, S.; Saitoh, E.; Sanada, K.

J. Biochem. 102, 693-704, 1987

A:Title: Characterization and amino acid sequence of a new acidic cysteine proteinase in

A:Reference number: A41422; MUID:88139220; PMID:3436950

A:Accession: A41422

A:Molecule type: protein

A:Residues: 25-141 <IS2>

C:Genetics:

A:Gene: GDB:GST2

A:Cross-references: GDB:119816; OMIM:123856

A:Map position: 20p11.2-20p11.2

C:Superfamily: cystatin; cystatin homology

F:30-141/Domain: cystatin homology <CYS>

Query Match 25.1%; Score 160.5; DB 2; Length 141;

Best Local Similarity 31.5%; Pred. No. 7.3e-09;

Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC--PETTNCVP 83

Db 47 ALHFVISEYNKATEDEYRLLRLVLRAREQIVGVNYPFDIEVGRITCTKSPNLDTCAP 106

QY 84 QER-ELHKQVCFPSVFAVPWFPEQYKILNKSC 114

Db 107 HQPELQKQLCSFQIYFVPEWDRMSLVNSRC 138

RESULT 6

A36163

cystatin C precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 16-Jul-1999

C:Accession: A36163

R:Solem, M.; Rawson, C.; Lindburg, K.; Barnes, D.

Biochem. Biophys. Res. Commun. 172, 945-951, 1990

A:Title: Transforming growth factor beta regulates cystatin C in serum-free mouse embryo

A:Reference number: A36163; MUID:91054522; PMID:2241983

A:Accession: A36163

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <SOL>

A:Cross-references: EMBL:M59470; NID:G192911; PIDN:AAA63298.1; PID:G192912

C:Superfamily: cystatin; cystatin homology

F:29-140/Domain: cystatin homology <CYS>

F:93-103,117-137/Disulfide bonds: #status predicted

Query Match 24.5%; Score 156.5; DB 2; Length 140;

Best Local Similarity 31.4%; Pred. No. 1.8e-08;

Matches 33; Conservative 24; Mismatches 45; Indels 3; Gaps 2;

QY 15 EYMAVENYAKDSLOWITDQYNKSDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC 74

Db 35 EADANEQVRRALDFAVSEYKNSGNDAYHSRAIQVVRARQLVAGVNYFDFVEMGRITCT 94

QY 75 KPET--TNC-VFQERELHKQVNCFFSVFAVPWFPEQYKILNKSCSS 116

Db 95 KSQTLNLTDCPFDDQPHLMRKALCSFQIYSPVWKGTHSLTKFSCKN 139

RESULT 7

UDRH

cystatin C precursor [validated] - human

N:Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma I

C:Species: Homo sapiens (man)

C>Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000

C:Accession: S10216; S00004; JLO095; A33400; S02751; A01270; A25434; S12288; A32732; A60

R:Abrahamson, M.; Olafsson, I.; Palsdottir, A.; Ulvsbaeck, M.; Lundwall, A.; Jensen, O.

Biochem. J. 268, 287-294, 1990

A:Title: Structure and expression of the human cystatin C gene.

A:Reference number: S10216; MUID:90303202; PMID:2363674

A:Accession: S10216

A:Molecule type: DNA

A:Residues: 1-146 <AB1>

R:Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FEBS Lett. 216, 229-233, 1987

A:Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A:Reference number: S00004; MUID:87219149; PMID:3495457

A:Accession: S00004

A:Molecule type: mRNA

A:Residues: 1-146 <AB2>

R:Levy, E.; Lopez-Otin, C.; Chiso, J.; Geltner, D.; Frangione, B.

J. Exp. Med. 169, 1771-1778, 1989

A:Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A:Reference number: JLO095; MUID:89235594; PMID:2541223

A:Accession: JLO095

A:Molecule type: DNA

A:Residues: 1-146 <LEV>

A:Cross-references: GB:X61681; NID:G30367; PIDN:CAA43856.2; PID:G4490944

A>Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit

e)

R:Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, K.

Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989

A:Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which i

A:Reference number: A33400; MUID:89350949; PMID:2764935

A:Accession: A33400

A:Molecule type: DNA

A:Residues: 1-24, 'T', 25-146 <SAI>

R:Chiso, J.; Cowan, N.; Frangione, B.

Biol. Chem. Hoppe-Seyler 369, 205-208, 1988

A:Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron

A:Reference number: S02751; MUID:89076507; PMID:3264504

A:Accession: S02751

A:Molecule type: DNA

A:Residues: 82-119 <GH2>

A:Cross-references: EMBL:M27769

A>Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown i

R:Grubb, A.; Lofberg, H.

Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982

A:Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the

A:Reference number: A01270; MUID:82222368; PMID:6283552

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 22.5179 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-4
Perfect score: 640
Sequence: 1 LPQARKKTFLSVHEVWAVE.....VFAVPWFQYKILNKSCSSD 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214	33.4	139	2 A45361	cystatin-related e
2	171.5	26.8	127	2 S07085	cystatin C precurs
3	170.5	26.6	139	1 UDCH	cystatin precursor
4	167.5	26.2	120	2 S10587	cystatin C - rat
5	160.5	25.1	141	2 B29632	cystatin SA precu
6	156.5	24.5	140	2 A36163	cystatin C precurs
7	155	24.2	146	1 UDHU	cystatin C precurs
8	153.5	24.0	112	1 UDBO	cystatin - bovine
9	146.5	22.9	141	1 UDHUP1	cystatin S precurs
10	139.5	21.8	141	1 UDHUP2	cystatin SN precu
11	134	20.9	111	1 JC2040	cystatin - chum sa
12	131.5	20.5	142	2 A47142	cystatin D precurs
13	129.5	20.2	141	2 JQ1470	cystatin S precurs
14	127	19.8	111	2 A28793	cystatin - puff ad
15	126	19.7	132	2 JC4918	cystatin precursor
16	102.5	16.0	133	2 JC4536	cystatin precursor
17	98.5	15.4	434	1 KGBOL2	kininogen, LMW II
18	98.5	15.4	619	1 KGBOH2	kininogen, HMW II
19	96.5	15.1	436	1 KGBOL1	kininogen, LMW I p
20	96.5	15.1	621	1 KGBOH1	kininogen, HMW I p
21	95.5	14.9	162	2 A43428	onchocystatin - ne
22	93.5	14.6	427	1 KGHUL1	kininogen, LMW pre
23	93.5	14.6	644	1 KGHUL1	kininogen, HMW pre
24	87	13.6	433	2 A28055	K-kininogen, LMW I
25	87	13.6	639	2 A28486	kininogen, HMW I p
26	85.5	13.4	430	2 A23897	major acute phase
27	81	12.7	430	2 B28055	T-kininogen, LMW I
28	80	12.5	498	2 T31871	hypothetical prote
29	79	12.3	430	1 KGRTT1	T-kininogen I prec

30	77	12.0	430	2 D64151	hypothetical prote
31	75	11.7	400	2 A46297	beta-1,6-N-acetylgl
32	75	11.7	423	1 KGRTM	major acute phase
33	74	11.6	582	2 S42613	membrane protein p
34	73.5	11.5	438	2 A47702	glucan 1,3-beta-gl
35	73.5	11.5	438	2 T52149	beta-glucanase lim
36	73	11.4	861	2 S12499	CHLI protein - yea
37	72.5	11.3	386	2 S14570	tubulin beta chain
38	71.5	11.2	132	2 D88508	protein H14A12.5 l
39	71.5	11.2	1227	2 T23004	hypothetical prote
40	71.5	11.2	1779	2 T23130	hypothetical prote
41	71	11.1	677	1 SYECMT	methionine-tRNA sy
42	71	11.1	677	2 C85839	methionine tRNA sy
43	71	11.1	677	2 H90993	methionine tRNA sy
44	70.5	11.0	448	2 JN0118	glucan 1,3-beta-gl
45	70.5	11.0	464	2 F81278	probable periplasm

ALIGNMENTS

RESULT 1

A45361
cystatin-related epididymal specific protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A45361
R;Cornwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.
Mol. Endocrinol. 6, 1653-1664, 1992
A;Title: The CRBS gene: a unique testis-regulated gene related to the cystatin family is
A;Reference number: A45361, MUID:93078799; PMID:1280328
A;Accession: A45361
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-139 <OR>
A;Cross-references: GB:S49926; NID:G260492; PIDN:AAC35390.1; PID:G260493
A;Note: sequence extracted from NCBI backbone (NCBIP:118813)
C;Superfamily: cystatin; cystatin homology
F;28-139/Domain: cystatin homology <Cys>

Query Match 33.4%; Score 214; DB 2; Length 139;
Best Local Similarity 37.1%; Pred. No. 3.5e-14;
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;

QY	12	SVHEWAVENY-----AKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLE	61
DB	22	SKNEVKA-QNYFGSINISNANVKQCVWFAMKYNKESDKYFLVDKILHAKLQITDRME	80
QY	62	YHLNVEMQWTTCKP--ETTNCVPQBR-ELHKQVNCFFSVFAVPWFQYKILNKSC	114
DB	81	YQIDVQISRNSCKPLNNTENCIPQKPLEKKNKSCFLVGALPWNGEFNLLSKEC	136

RESULT 2

S07085

cystatin C precursor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999
C;Accession: S07085; S01337; S21109
R;Cole, T.; Dickson, P.W.; Esnard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Schre
Eur. J. Biochem. 186, 35-42, 1989
A;Title: The cDNA structure and expression analysis of the genes for the cysteine protein
A;Reference number: S07085; MUID:90092122; PMID:2689174
A;Accession: S07085
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-127 <COL>
A;Cross-references: EMBL:X16957; NID:G56041; PIDN:CAA34831.1; PID:G736290
R;Esnard, A.; Esnard, F.; Faucher, D.; Gauthier, F.
FEBS Lett. 236, 475-478, 1988
A;Title: Two rat homologues of human cystatin C.
A;Reference number: S01337; MUID:88313020; PMID:3044831
A;Accession: S01337

; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-6

Query Match 29.5%; Score 189; DB 9; Length 36;
 Best Local Similarity 100.0%; Pred. No. 4.1e-14;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKTFLSVHVEVMAVENYAKDSLOWITDQYNKESDDK 41
 Db 1 RKTFLSVHVEVMAVENYAKDSLOWITDQYNKESDDK 36
 RESULT 14
 US-09-941-314-8
 ; Sequence 8, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 35
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-8

Query Match 29.2%; Score 187; DB 9; Length 35;
 Best Local Similarity 100.0%; Pred. No. 6.7e-14;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 58
 Db 1 KDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTD 35
 RESULT 15
 US-09-740-638-2
 ; Sequence 2, Application US/09740638
 ; Patent No. US20020006656A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Holloway, James L.
 ; TITLE OF INVENTION: Zcys85: A Member of the Cystatin
 ; TITLE OF INVENTION: Superfamily
 ; FILE REFERENCE: 99-104
 ; CURRENT APPLICATION NUMBER: US/09/740,638
 ; CURRENT FILING DATE: 2000-12-18
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Fast-SEQ for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 145
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-740-638-2

Query Match 27.7%; Score 177.5; DB 9; Length 145;

Best Local Similarity 31.4%; Pred. No. 4.5e-12;
 Matches 32; Conservative 30; Mismatches 37; Indels 3; Gaps 2;
 QY 15 EVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVEHQWITCQ 74
 Db 34 KLMSKKN-MNSTLNFPIQSYNNASNDTYLYRVORLIIRSQRLTGTGVEYIVTVKIGWTCKK 92
 QY 75 KPETTIN--CVPQERELHKOVCFFSVFAVPWFQYKILNKSC 114
 Db 93 RNDTSNSSCPLOSKKLKRLKSLICESLIYTMPEWYFQLWNNSC 134

Search completed: March 18, 2004, 14:30:52
 Job time : 70.8036 secs

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKKTLFSLVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55
 Db 1 RKKTLFSLVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 10
 US-09-864-761-48936
 ; Sequence 48936, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 48936
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL096677.18
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
 ; OTHER INFORMATION: EST HUMAN HIT: AI200857.1, EVALU6 5.00e-23
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALU6 1.00e-01
 US-09-864-761-48936

Query Match 40.98; Score 262; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 3.4e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKKTLFSLVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55
 Db 1 RKKTLFSLVHVMVAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 11
 US-09-941-314-10
 ; Sequence 10, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR APPLICATION NUMBER: 60/230,230
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 46
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-10

Query Match 39.7%; Score 254; DB 9; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e-21;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 80
 Db 1 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 46

RESULT 12
 US-09-941-314-12
 ; Sequence 12, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR APPLICATION NUMBER: 60/230,230
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 33
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-12

Query Match 29.5%; Score 189; DB 9; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.7e-14;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 QVTDHLEYHLNVEMQWTTCKPPTTNVCVPOERE 87
 Db 1 QVTDHLEYHLNVEMQWTTCKPPTTNVCVPOERE 33

RESULT 13
 US-09-941-314-6
 ; Sequence 6, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to

```

; ORGANISM: Homo sapiens
US-09-941-314-14
Query Match 45.0%; Score 288; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 4.1e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KESDDKYHRIPIRVLVKQVQVTDHLEVLHNVEMQWTTCKPPTNCVPOERE 87
Db 1 KESDDKYHRIPIRVLVKQVQVTDHLEVLHNVEMQWTTCKPPTNCVPOERE 52

RESULT 7
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-17

Query Match 42.7%; Score 273; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 NVEMQWTTCKPPTNCVPOERLHKQVNCFFSVFVPMFEQYKILNK 112
Db 1 NVEMQWTTCKPPTNCVPOERLHKQVNCFFSVFVPMFEQYKILNK 48

RESULT 8
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-13

Query Match 42.5%; Score 272; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DQYNKESDDKYHRIPIRVLVKQVQVTDHLEVLHNVEMQWTTCKPPTN 80
Db 1 DQYNKESDDKYHRIPIRVLVKQVQVTDHLEVLHNVEMQWTTCKPPTN 49

RESULT 9
US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34822
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109954.10
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match 40.9%; Score 262; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;

```

```
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; PRIOR FILING DATE: 2001-08-29
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-2

Query Match          100.0%; Score 640; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPYQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHL 60
DB 21 LPYQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHL 80

QY 61 EYHLNVEMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFPEQYKILNKSCSSD 117
DB 81 EYHLNVEMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFPEQYKILNKSCSSD 137

RESULT 3
US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-3

Query Match          98.3%; Score 629; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 3e-63;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEY 62
DB 1 YQARKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEY 60

QY 63 HLNVEQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFPEQYKILNKSCSSD 117
DB 61 HLNVEQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFPEQYKILNKSCSSD 115

RESULT 4
US-09-941-314-15
; Sequence 15, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
```

```
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-15

Query Match          69.7%; Score 446; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 9.6e-43;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 QYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCKPETTNCVQRELHKQV 92
DB 1 QYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCKPETTNCVQRELHKQV 60

QY 93 NCFPSVFAVPWFPEQYKILNK 112
DB 61 NCFPSVFAVPWFPEQYKILNK 80

RESULT 5
US-09-941-314-16
; Sequence 16, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-16

Query Match          52.2%; Score 334; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 3e-30;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 RQVTDHLEYHLNVEMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFPEQYKILNK 112
DB 1 RQVTDHLEYHLNVEMQWTTCKPETTNCVQRELHKQVNCFFSVFVAVPWFPEQYKILNK 59

RESULT 6
US-09-941-314-14
; Sequence 14, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 70.8036 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-4
Perfect score: 640
Sequence: 1 LPQARKKTFSLVHEWAVE.....VFAVPWFQYKILNKSCSSD 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap:
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap:
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap:
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap:
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap:
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap:
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap:
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	640	100.0	117	9	US-09-941-314-4 Sequence 4, Appli
2	640	100.0	137	9	US-09-941-314-2 Sequence 2, Appli
3	629	98.3	115	9	US-09-941-314-3 Sequence 3, Appli
4	446	69.7	80	9	US-09-941-314-15 Sequence 15, Appl
5	334	52.2	59	9	US-09-941-314-16 Sequence 16, Appl
6	288	45.0	52	9	US-09-941-314-14 Sequence 14, Appl
7	273	42.7	48	9	US-09-941-314-17 Sequence 17, Appl
8	272	42.5	49	9	US-09-941-314-13 Sequence 13, Appl
9	262	40.9	50	9	US-09-864-761-34822 Sequence 14822, A
10	262	40.9	50	9	US-09-864-761-48936 Sequence 48936, A
11	254	39.7	46	9	US-09-941-314-10 Sequence 10, Appl
12	189	29.5	33	9	US-09-941-314-12 Sequence 12, Appl
13	189	29.5	36	9	US-09-941-314-6 Sequence 6, Appli
14	187	29.2	35	9	US-09-941-314-8 Sequence 8, Appli
15	177.5	27.7	145	9	US-09-740-638-2 Sequence 2, Appli

16	177.5	27.7	145	13	US-10-006-467-2	Sequence 2, Appli
17	177.5	27.7	145	14	US-10-235-148-2	Sequence 2, Appli
18	171.5	26.8	127	8	US-08-849-303-19	Sequence 19, Appl
19	171.5	26.8	145	14	US-10-168-425-14	Sequence 14, Appl
20	170.5	26.6	116	9	US-09-775-932-16	Sequence 16, Appl
21	170.5	26.6	139	8	US-08-849-303-15	Sequence 15, Appl
22	170.5	26.6	139	9	US-09-969-834-4	Sequence 4, Appli
23	160.5	25.1	121	9	US-09-775-932-8	Sequence 8, Appli
24	160.5	25.1	140	14	US-10-376-564-46	Sequence 46, Appl
25	160.5	25.1	141	8	US-08-849-303-24	Sequence 24, Appl
26	160.5	25.1	141	9	US-09-940-497-6	Sequence 6, Appli
27	159.5	24.9	181	15	US-10-264-049-2608	Sequence 2608, Ap
28	157.5	24.6	165	9	US-09-740-638-5	Sequence 5, Appli
29	157.5	24.6	165	13	US-10-006-467-5	Sequence 5, Appli
30	157.5	24.6	165	14	US-10-235-148-5	Sequence 18, Appl
31	156.5	24.5	140	8	US-08-849-303-18	Sequence 2, Appli
32	156.5	24.5	140	14	US-10-376-564-48	Sequence 48, Appl
33	155	24.2	120	9	US-09-775-932-2	Sequence 17, Appl
34	155	24.2	146	8	US-08-849-303-17	Sequence 3, Appli
35	155	24.2	146	9	US-09-940-497-3	Sequence 3, Appli
36	155	24.2	146	9	US-09-969-834-3	Sequence 3, Appli
37	155	24.2	146	14	US-10-329-428-3	Sequence 47, Appl
38	155	24.2	146	14	US-10-376-564-47	Sequence 16, Appl
39	153.5	24.0	112	8	US-08-849-303-16	Sequence 24, Appl
40	153.5	24.0	118	9	US-09-775-932-24	Sequence 4, Appli
41	146.5	22.9	121	9	US-09-775-932-4	Sequence 22, Appl
42	146.5	22.9	141	8	US-08-849-303-22	Sequence 5, Appli
43	146.5	22.9	141	9	US-09-940-497-5	Sequence 141, App
44	146.5	22.9	141	9	US-09-974-298-141	Sequence 77, Appl
45	146.5	22.9	141	14	US-10-241-220-77	

ALIGNMENTS

RESULT 1
US-09-941-314-4
; Sequence 4, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-4

Query Match 100.0%; Score 640; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.8e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LPQARKKTFSLVHEWAVENYAKDSIQWITDQYNKESDDKYHFRIVRLKQVQVTDHL 60
Db 1 LPQARKKTFSLVHEWAVENYAKDSIQWITDQYNKESDDKYHFRIVRLKQVQVTDHL 60
Qy 61 EYHNVEMQWTTQCKPETTNCVPERELHKQVNCFFSVFVAPWFPEQYKILNKSCSSD 117
Db 61 EYHNVEMQWTTQCKPETTNCVPERELHKQVNCFFSVFVAPWFPEQYKILNKSCSSD 117
RESULT 2
US-09-941-314-2
; Sequence 2, Application US/09941314
; Patent No. US20020142396A1


```
; TELEFAX: 301 309 8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Cystatin SN
; US-08-744-138-6

Query Match      25.1% Score 160.5; DB 3; Length 141;
Best Local Similarity 31.5%; Pred. No. 5.1e-12;
Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY      26 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEHYHLNVMQWTTCK--PETTNCVP 83
Db      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
47 ALHFVISEYNKATEDEYRRLRLRLRAREQIVGGVNYFFDIEVGRITCTKSPNLDTCF 106

QY      84 QER-ELHKQVNCFFSVFAVPWFQYKILNKSC 114
Db      |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
107 HEQPELQKKQLCSFQIYEVPWEDRMSLVNSRC 138
```

Search completed: March 18, 2004, 14:25:34
Job time : 29.9464 secs

Db 107 HDPEMAKYTTCTFFVYVSWPWLQIKLLESKC 138

RESULT 12

5432264-6

Patent No. 5432264

APPLICANT: GRUBB, ANDERS; LUNDWALL, AKE; ABRAHAMSON, MAGNUS;

DALBOGE, HENRIK

TITLE OF INVENTION: RECOMBINANT 3-DES-OH-CYSTATIN C PRODUCED

BY EXPRESSION IN A PROCARYOTIC HOST CELL

NUMBER OF SEQUENCES: 8

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/929,290

FILING DATE: 13-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 440,221

FILING DATE: 21-NOV-1989

APPLICATION NUMBER: 297,198

FILING DATE: 20-MAY-1988

SEQ ID NO: 6;

LENGTH: 146

5432264-6

Query Match 25.3%; Score 162; DB 6; Length 146;

Best Local Similarity 34.7%; Pred. No. 3.5e-12;

Matches 35; Conservative 21; Mismatches 41; Indels 4; Gaps 3;

QY 20 ENYAKDSLOWITDOYNKESDDKYHFRIVRLKQVQVTDHLEYHLNVEMQWTTQK--PE 77

Db 46 EEGVRRLDPFVGEYNKASNDYHSRALQVVRARKQIVAGVNYFLDVLGRTTCTKTQPN 105

QY 78 TTNC-VPOERELHKQVNCFFSVFVAPVFEQYKILNKSCSD 117

Db 106 LDNCPHFDQPHLRKAFCSQIYVAPW-QQTWTLSTKSTOD 145

RESULT 13

US-09-775-932-8

Sequence 8, Application US/09775932

Patent No. 6534477

GENERAL INFORMATION:

APPLICANT: University of British Columbia

TITLE OF INVENTION: Production and use of Modified Cystatins

FILE REFERENCE: 58069

CURRENT APPLICATION NUMBER: US/09/775,932

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: CA99/00717

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 60/095,503

PRIOR FILING DATE: 1998-08-05

NUMBER OF SEQ ID NOS: 32

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 8

LENGTH: 121

TYPE: PRT

ORGANISM: Homo sapiens

US-09-775-932-8

Query Match 25.1%; Score 160.5; DB 4; Length 121;

Best Local Similarity 31.5%; Pred. No. 4.2e-12;

Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDOYNKESDDKYHFRIVRLKQVQVTDHLEYHLNVEMQWTTQK--PETTNCVP 83

Db 27 ALHPVISEYNKATEDEYRLLAVLRAREQIVGVNYFFDIEVGRITCTKSQPNLDTCAF 86

QY 84 QER-ELHKQVNCFFSVFVAPVFEQYKILNKSC 114

Db 87 HEQPELQKQLCSFQIYVEVPEWEDRMSLVNSRC 118

RESULT 14

US-09-886-319A-46

Sequence 46, Application US/09886319A

Patent No. 6596185

GENERAL INFORMATION:

APPLICANT: Wolf, Eckard

APPLICANT: Werner, Sabine

APPLICANT: Halle, Jorn-Peter

APPLICANT: Regenbogen, Johannes

APPLICANT: Goppelt, Andreas

TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for

TITLE OF INVENTION: the diagnosis or Treatment of Skin Disorders and Wound

TITLE OF INVENTION: Healing and for the Identification of Pharmacologically

TITLE OF INVENTION: Active Substances

FILE REFERENCE: 50125/014002

CURRENT APPLICATION NUMBER: US/09/886,319A

CURRENT FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: US 60/222,081

PRIOR FILING DATE: 2000-08-01

PRIOR APPLICATION NUMBER: DE 10030149.5

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 84

SOFTWARE: PastSeq for Windows Version 4.0

SEQ ID NO 46

LENGTH: 140

TYPE: PRT

ORGANISM: Mus musculus

US-09-886-319A-46

Query Match 25.1%; Score 160.5; DB 4; Length 140;

Best Local Similarity 32.4%; Pred. No. 5.1e-12;

Matches 34; Conservative 24; Mismatches 44; Indels 3; Gaps 2;

QY 15 EYMAVENYAKDSLOWITDOYNKESDDKYHFRIVRLKQVQVTDHLEYHLNVEMQWTTQ 74

Db 35 EADANEEGVRRALDFAVSEYNGKSDAYHSRAIQVVRARKQIVAGVNYFLDVEMGRTTCT 94

QY 75 KPET--TNC-VPOERELHKQVNCFFSVFVAPVFEQYKILNKSCSS 116

Db 95 KSQTNLTDCPFHDQPHLMRKALCSFQIYSPVPMKGTSLTKFSCKN 139

RESULT 15

US-08-744-138-6

Sequence 6, Application US/08744138

Patent No. 6011012

GENERAL INFORMATION:

APPLICANT: Gentz, Reiner L.

APPLICANT: Ni, Jian

APPLICANT: Rosen, Craig A.

APPLICANT: Yu, Guo-Liang

TITLE OF INVENTION: Human Cystatin E

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/744,138

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF202P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301 309 8504

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
;
US-08-791-522-4
Query Match 26.6%; Score 170.5; DB 2; Length 139;
Best Local Similarity 37.0%; Pred. No. 3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFRLVKVQRQVTDHLEYHLNVEMQWTTCKP--ETNC-V 82
Db 47 ALQFAMAEYNRASNDKYSSRVVRVISAKRLQVSGIKYILOVEIGRTTCPKSSGDLOSCEF 106

QY 83 POERELHKQVNCFFSVFVFWFEQYKILNKSC 114
Db 107 HDEPEMAKYTTCTFVVVYSIPWLNQIKLLESKC 138

RESULT 10
US-09-314-777-4
; Sequence 4, Application US/09314777
; Patent No. 6110686
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Sufya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,777
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,522
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
;
US-09-314-777-4
Query Match 26.6%; Score 170.5; DB 3; Length 139;
Best Local Similarity 37.0%; Pred. No. 3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFRLVKVQRQVTDHLEYHLNVEMQWTTCKP--ETNC-V 82
Db 47 ALQFAMAEYNRASNDKYSSRVVRVISAKRLQVSGIKYILOVEIGRTTCPKSSGDLOSCEF 106

QY 83 POERELHKQVNCFFSVFVFWFEQYKILNKSC 114
Db 107 HDEPEMAKYTTCTFVVVYSIPWLNQIKLLESKC 138

RESULT 11
US-08-849-303-15
; Sequence 15, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urwin, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
;
INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
;
US-08-849-303-15
Query Match 26.6%; Score 170.5; DB 4; Length 139;
Best Local Similarity 37.0%; Pred. No. 3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLOWITDQYNKSDDKYHFRIFRLVKVQRQVTDHLEYHLNVEMQWTTCKP--ETNC-V 82
Db 47 ALQFAMAEYNRASNDKYSSRVVRVISAKRLQVSGIKYILOVEIGRTTCPKSSGDLOSCEF 106

QY 83 POERELHKQVNCFFSVFVFWFEQYKILNKSC 114
```

```

; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-4

Query Match      29.5%; Score 188.5; DB 3; Length 142;
Best Local Similarity 40.4%; Pred. No. 1.9e-15;
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;

QY 24 KDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCCQKPTTN--C 81
Db 46 KQCLWFAMQYNEKSEDDKYVFLVVKTLQALQVTLNLEYLIDVIEIARSDCRKPLSTNEIC 105

QY 82 VPQER-ELHKQVNCFFSVFVAVPWFPEQYKILNKSC 114
Db 106 AIGENSKRLKSLCSFLVGLPWNGETVMEKCC 139

RESULT 7
US-08-849-303-19
; Sequence 19, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urwin, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-849-303-19

Query Match      26.8%; Score 171.5; DB 4; Length 127;
Best Local Similarity 33.3%; Pred. No. 2e-13;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

```

```

QY 15 EYMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCC 74
Db 22 EADASEGVQRALDFAVSEYNGKSGNDAYHSRAIQVVRARQQLVAGINYLYLDVEMGRITCT 81

QY 75 KPET--TNC-VPQERELHKQVNCFFSVFVAVPWFPEQYKILNKSCSS 116
Db 82 KSQTNLTNCPFFHDQPHLMKALCSFQIYSVPWKGTHTLTAKSSCKN 126

RESULT 8
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cystatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match      26.6%; Score 170.5; DB 4; Length 116;
Best Local Similarity 37.0%; Pred. No. 2.3e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 26 SLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCCQK--ETTNC-V 82
Db 24 ALQFAMAEYNRASNDKYSRVRVISAKQLVSGIKYILQVEIGRTTCPKSSGDLQSCF 83

QY 83 PQERELHKQVNCFFSVFVAVPWFPEQYKILNKSC 114
Db 84 HDEPEMAKYTTCTFVVYVIPNLNQLKLESKC 115

RESULT 9
US-08-791-522-4
; Sequence 4, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,522
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-617-302-3

Query Match 33.4%; Score 214; DB 3; Length 142;
Best Local Similarity 37.4%; Pred. No. 1.4e-18;
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;
QY 12 SVHEVMAVENY-----AKDSLOWITDOYNKESDDKYHFRIPRLVKVQVQVTDHLE 61
DB 25 SKNEVA-QNYFSGINISNANVCQWPFAMKEYNKESDKYFLVDKILHAKLQITDRME 83
QY 62 YHLNVMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114
DB 84 YQIDVOISRSNCKPLNNTENCIPQKPELEKQMSCSFLVGALPWNGEFNLKSEC 139

RESULT 3
US-09-431-480-2
; Sequence 2, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-2

Query Match 29.5%; Score 188.5; DB 3; Length 141;
Best Local Similarity 34.5%; Pred. No. 1.9e-15;
Matches 40; Conservative 25; Mismatches 38; Indels 13; Gaps 3;
QY 2 PYQARKKTFLSVHEVMAVENYAKDSLOWITDOYNKESDDKYHFRIPRLVKVQVQVTDHLE 61
DB 33 PFEDIPKSVY-----YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLE 82
QY 62 YHLNVMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114
DB 83 YYLEVNIARTMCKKIAGDNENCLFQDDPKMKWVFCIFIVSSKPKWPKELKMLKKQC 138

RESULT 4
US-09-617-302-2
; Sequence 2, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-2

Query Match 29.5%; Score 188.5; DB 3; Length 141;
Best Local Similarity 34.5%; Pred. No. 1.9e-15;
Matches 40; Conservative 25; Mismatches 38; Indels 13; Gaps 3;
QY 2 PYQARKKTFLSVHEVMAVENYAKDSLOWITDOYNKESDDKYHFRIPRLVKVQVQVTDHLE 61
DB 33 PFEDIPKSVY-----YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLE 82
QY 62 YHLNVMQWTTCKP--ETTNCVPQER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114
DB 83 YYLEVNIARTMCKKIAGDNENCLFQDDPKMKWVFCIFIVSSKPKWPKELKMLKKQC 138

RESULT 5
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match 29.5%; Score 188.5; DB 3; Length 142;
Best Local Similarity 40.4%; Pred. No. 1.9e-15;
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;
QY 24 KDSLOWITDOYNKESDDKYHFRIPRLVKVQVQVTDHLEHLNVMQWTTCKPPTTN--C 81
DB 46 KQCLWFAMQYKNESEDKYFLVVKTLQALQVTLNLELYLIDVEIARSDCKPLSTNEIC 105
QY 82 VPOER-ELHKQVNCFFSVFVFPWFPEQYKILNKSC 114
DB 106 AIQENSKKLKLSCSFLVGALPWNGEFTVMEKCC 139

RESULT 6
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR FILING DATE: 09/431,480

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 29.9464 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-4
Perfect score: 640
Sequence: 1 LPYQARKTKYFLSVHEWAVE.....VFAPWFEQYKILNKSCSD 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	33.4	142	3	US-09-431-480-3
2	214	33.4	142	3	US-09-617-302-3
3	188.5	29.5	141	3	US-09-431-480-2
4	188.5	29.5	141	3	US-09-617-302-2
5	188.5	29.5	142	3	US-09-431-480-4
6	188.5	29.5	142	3	US-09-617-302-4
7	171.5	26.6	127	4	US-08-849-303-19
8	170.5	26.6	116	4	US-09-775-932-16
9	170.5	26.6	139	2	US-08-791-522-4
10	170.5	26.6	139	3	US-09-314-777-4
11	170.5	26.6	139	4	US-08-849-303-15
12	162	25.3	146	6	5432264-6
13	160.5	25.1	121	4	US-09-775-932-8
14	160.5	25.1	140	4	US-09-886-319A-46
15	160.5	25.1	141	3	US-08-744-138-6
16	160.5	25.1	141	4	US-09-241-376-6
17	160.5	25.1	141	4	US-09-940-497-6
18	160.5	25.1	141	4	US-08-849-303-24
19	156.5	24.5	140	3	US-09-431-480-5
20	156.5	24.5	140	3	US-09-617-302-5
21	156.5	24.5	140	4	US-09-886-319A-48
22	156.5	24.5	140	4	US-08-849-303-18
23	155	24.2	120	4	US-09-775-932-2
24	155	24.2	120	6	5432264-4
25	155	24.2	145	2	US-08-832-535-11
26	155	24.2	146	2	US-08-791-522-3
27	155	24.2	146	3	US-08-744-138-3

28	155	24.2	146	3	US-09-019-485-4	Sequence 4, Appli
29	155	24.2	146	3	US-09-314-777-3	Sequence 3, Appli
30	155	24.2	146	3	US-09-431-480-6	Sequence 6, Appli
31	155	24.2	146	3	US-09-617-302-6	Sequence 6, Appli
32	155	24.2	146	4	US-09-241-376-3	Sequence 3, Appli
33	155	24.2	146	4	US-09-528-436B-3	Sequence 3, Appli
34	155	24.2	146	4	US-09-886-319A-47	Sequence 47, Appli
35	155	24.2	146	4	US-09-940-497-3	Sequence 3, Appli
36	155	24.2	146	4	US-09-976-594-37	Sequence 37, Appli
37	155	24.2	146	4	US-08-849-303-17	Sequence 17, Appli
38	155	24.2	146	5	PCT-US95-07135-9	Sequence 9, Appli
39	153.5	24.0	112	4	US-08-849-303-16	Sequence 16, Appli
40	153.5	24.0	118	4	US-09-775-932-24	Sequence 24, Appli
41	146.5	22.9	121	4	US-09-775-932-4	Sequence 4, Appli
42	146.5	22.9	141	3	US-08-744-138-5	Sequence 5, Appli
43	146.5	22.9	141	3	US-09-431-480-11	Sequence 11, Appli
44	146.5	22.9	141	3	US-09-617-302-11	Sequence 11, Appli
45	146.5	22.9	141	4	US-09-241-376-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-431-480-3
; Sequence 3, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-480-3

Query Match	33.4%	Score 214;	DB 3;	Length 142;
Best Local Similarity	37.1%	Pred. NO. 1.4e-18;		
Matches	43;	Conservative 29;	Mismatches 30;	Indels 14; Gaps 4;
QY	12	SVHEWAVENY-----AKDSLOWITDQNKESDDKYHFRIFRVLKVRQVTDHLE	61	
DB	25	SKNEVKA-QNYFGSINISNANVKQCVFAMKEYNESDKYVFLVDKILHAKLQITDRME	83	
QY	62	YHLNVEWMTTCOKP--ETTNCVPOER-ELHKOVNCFESVFAVPWFEOYKILNKSC	114	
DB	84	YQIDVOISRNCKPLNNTENCIPQKKPELEKKNCSFLVGALPFWNGEFLNLSKEC	139	

RESULT 2
US-09-617-302-3
; Sequence 3, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts.
 PT
 PT
 PS Claim 15; SEQ ID NO 21294; 530pp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA411305). the present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 50 AA;
 SQ

Query Match 40.9%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 15
 AM67252
 ID AM67252 standard; protein; 50 AA.
 XX
 XX AM67252;
 AC
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.
 XX
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 XX Homo sapiens.
 OS
 XX WO200157276-A2.
 PN
 XX
 XX 09-AUG-2001.
 PD
 XX
 XX 30-JAN-2001; 2001WO-US000668.
 PF
 XX
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX
 XX WPI; 2001-488900/53.
 DR
 XX
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 PT
 XX
 XX Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.
 PS
 XX

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention
 XX
 XX Sequence 50 AA;
 SQ

Query Match 40.9%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 55
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

Search completed: March 18, 2004, 14:14:56
 Job time : 106.089 secs

XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-496933/54.
 DR New spatially-addressable set of single exon nucleic acid probes, useful
 XX for measuring gene expression in sample derived from human breast,
 XX PT comprises number of single exon nucleic acid probes.
 XX Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.
 PS The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a peptide
 CC encoded by a single exon nucleic acid probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 50 AA;
 SQ

Query Match 40.9%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RKKTFLSVHEVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 55
 DB 1 RKKTFLSVHEVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 50

RESULT 13
 ABB28913
 ID ABB28913 standard; peptide; 50 AA.
 XX ABB28913;
 XX 01-FEB-2002 (first entry)
 DT Peptide #1564 encoded by breast cell single exon nucleic acid probe.
 DE Human; microarray; single exon probe; gene expression; breast; disease;
 KW cancer.
 KW Homo sapiens.
 OS WO200157271-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000662.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-496933/54.
 DR New spatially-addressable set of single exon nucleic acid probes, useful
 XX for measuring gene expression in sample derived from human breast,
 XX PT comprises number of single exon nucleic acid probes.
 XX Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.
 PS The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a peptide
 CC encoded by a single exon nucleic acid probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 50 AA;
 SQ

Query Match 40.9%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RKKTFLSVHEVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 55
 DB 1 RKKTFLSVHEVWAVENYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQRQ 50

RESULT 14
 ABB19524
 ID ABB19524 standard; protein; 50 AA.
 XX ABB19524;
 XX 23-JAN-2002 (first entry)
 DT Protein #1523 encoded by probe for measuring heart cell gene expression.
 DE Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 KW Homo sapiens.
 OS WO200157274-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000666.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA

OS	Homo sapiens.
XX	WO200157272-A2.
XX	
XX	09-AUG-2001.
XX	
XX	30-JAN-2001; 2001WO-US000663.
XX	
XX	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.
XX	
XX	(MOLE-) MOLECULAR DYNAMICS INC.
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
XX	WPI; 2001-48897/S3.
XX	
DR	Human genome-derived single exon nucleic acid probes useful for analyzing
PT	gene expression in human placenta.
PT	
XX	
XX	Claim 27; SEQ ID NO 27814; 654pp; English.
XX	
CC	The present invention relates to single exon nucleic acid probes (SENP:
CC	see AA131315-AA157546). The present sequence is a peptide encoded by one
CC	such probe. The probes are useful for producing a microarray for
CC	predicting, measuring and displaying gene expression in samples derived
CC	from human placenta. The probes are useful for antenatal diagnosis of
CC	human genetic disorders
XX	
XX	Sequence 50 AA;
SQ	
	Query Match 40.9%; Score 262; DB 4; Length 50;
	Best Local Similarity 100.0%; Pred. No. 2.9e-22;
	Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	6 RKKTFLSVHEHVAVENYAKDSLQWITDQYNKESDDKYHFIRFVLKVQRQ 55
DB	1 RKKTFLSVHEHVAVENYAKDSLQWITDQYNKESDDKYHFIRFVLKVQRQ 50
RESULT 12	
ABB32389	
ID	ABB32389 standard; peptide; 50 AA.
XX	
XX	ABB32389;
XX	
DT	
DE	01-FEB-2002 (first entry)
XX	
XX	Peptide #5040 encoded by breast cell single exon nucleic acid probe.
KW	Human; microarray; single exon probe; gene expression; breast; disease;
KW	cancer.
XX	
OS	Homo sapiens.
XX	
XX	WO200157271-A2.
PN	
PD	
XX	09-AUG-2001.
XX	
XX	30-JAN-2001; 2001WO-US000662.
XX	
XX	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX SQ Sequence 48 AA;
 Query Match 42.7%; Score 273; DB 5; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.5e-23;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 NVEMQWTTCKPNTTNCVQERLHKQVNCFFSVFAVPWFQYKILNK 112
 |||||
 Db 1 NVEMQWTTCKPNTTNCVQERLHKQVNCFFSVFAVPWFQYKILNK 48

RESULT 8
 AAU79863
 ID AAU79863 standard; peptide; 49 AA.
 XX
 AC AAU79863;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #11.
 XX
 DE Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 23-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230320P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 XX Holloway JL, Gao Z, Bishop PD;
 XX WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 97-98; 100pp; English.
 XX

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX SQ Sequence 49 AA;
 Query Match 42.5%; Score 272; DB 5; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DOYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCKPETT 80
 |||||
 Db 1 DOYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCKPETT 49

RESULT 9
 AAM15096
 ID AAM15096 standard; protein; 50 AA.
 XX
 AC AAM15096;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #1530 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000670.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX
 PS Claim 27; SEQ ID NO 19922; 487pp; English.
 XX

CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX SQ Sequence 50 AA;
 Query Match 40.9%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RKTFTLSVHEWAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQV 55
 |||||
 Db 1 RKTFTLSVHEWAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQV 50

CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 59 AA;

Query Match 52.2%; Score 334; DB 5; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.2e-30; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 0;

QY 54 RQVTDHLEYHLNVEMQWTTCKPQRETLHKQVNCFFSFVAFWPFQYKILNK 112
Db 1 RQVTDHLEYHLNVEMQWTTCKPQRETLHKQVNCFFSFVAFWPFQYKILNK 59

RESULT 6
AAU79864
ID AAU79864 standard; peptide; 52 AA.

XX AAU79864;
XX
XX 15-JUL-2002 (first entry)
XX Human cystatin-8 (Zcys8) antigenic fragment #12.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
XX spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
XX sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.
XX WO200220567-A2.
XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.
XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;
XX WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 52 AA;

Query Match 45.0%; Score 288; DB 5; Length 52;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KESDDKYHFRIPRLVKVQRQVTDHLEYHLNVEMQWTTCKPQRETLHKQVNCFFSFVAFWPFQYKILNK 87
Db 1 KESDDKYHFRIPRLVKVQRQVTDHLEYHLNVEMQWTTCKPQRETLHKQVNCFFSFVAFWPFQYKILNK 52

RESULT 7
AAU79867
ID AAU79867 standard; peptide; 48 AA.

XX AAU79867;

XX 15-JUL-2002 (first entry)

XX Human cystatin-8 (Zcys8) antigenic fragment #15.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
XX spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
XX sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 99; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 115 AA;

Query Match 98.3%; Score 629; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1.2e-63;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YQARKKTFSLVHVMVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 62
 DB 1 YQARKKTFSLVHVMVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 60
 QY 63 HLNVENQWTTCKPPTTNCVQPERELHKQVNCFFSVFAVPWFQYKILNKSCSSD 117
 DB 61 HLNVENQWTTCKPPTTNCVQPERELHKQVNCFFSVFAVPWFQYKILNKSCSSD 115

RESULT 4
 AAU79865
 ID AAU79865 standard; peptide; 80 AA.
 AC AAU79865;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #13.
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 98; 100pp; English.

XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 80 AA;

Query Match 69.7%; Score 446; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 5.6e-43;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 QYNKESDDKYHFRIFRLVKVQROVTDHLEYHLNVENQWTTCKPPTTNCVQPERELHKQV 92
 DB 1 QYNKESDDKYHFRIFRLVKVQROVTDHLEYHLNVENQWTTCKPPTTNCVQPERELHKQV 60
 QY 93 NCFPSVFAVPWFQYKILNK 112
 DB 61 NCFPSVFAVPWFQYKILNK 80

RESULT 5
 AAU79866
 ID AAU79866 standard; peptide; 59 AA.
 AC AAU79866;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #14.
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 99; 100pp; English.

XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC polynucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 CC
 XX
 SQ Sequence 117 AA;

Query Match 100.0%; Score 640; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 7e-65;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 60
 DB 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 60
 QY 61 EYHLNVMQWTTCKPETTNCVQPERLHKQVNCFFSVFAPWPFQYKILNKSCSSD 117
 DB 61 EYHLNVMQWTTCKPETTNCVQPERLHKQVNCFFSVFAPWPFQYKILNKSCSSD 117

RESULT 2
 AAU79852
 ID AAU79852 standard; protein; 137 AA.
 XX
 AC AAU79852;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8).
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 DR N-PSDB; ABK49522.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 93-94; 100pp; English.
 XX

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like

CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 137 AA;

Query Match 100.0%; Score 640; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 8.6e-65;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 60
 DB 21 LPYQARKKTFSLVHEVMAVENYAKDSLQWITDOYNKESDDKYHFRIFRVLKQVQVTDHL 80
 QY 61 EYHLNVMQWTTCKPETTNCVQPERLHKQVNCFFSVFAPWPFQYKILNKSCSSD 117
 DB 81 EYHLNVMQWTTCKPETTNCVQPERLHKQVNCFFSVFAPWPFQYKILNKSCSSD 137

RESULT 3
 AAU79853
 ID AAU79853 standard; protein; 115 AA.
 XX
 AC AAU79853;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #1.
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 94; 100pp; English.
 XX

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 106.089 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-4
Perfect score: 640
Sequence: 1 LPQARKTKFLSVHEWAVE.....VFAVPWFQYKILNKSCSSD 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep_29Jan04:.*
1: Genesep1980s:.*
2: Genesep1990s:.*
3: Genesep2000s:.*
4: Genesep2001s:.*
5: Genesep2002s:.*
6: Genesep2003as:.*
7: Genesep2003bs:.*
8: Genesep2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	640	100.0	117	5	AAU79854 Human cys
2	640	100.0	137	5	AAU79852 Human cys
3	629	98.3	115	5	AAU79853 Human cys
4	446	69.7	80	5	AAU79865 Human cys
5	334	52.2	59	5	AAU79866 Human cys
6	288	45.0	52	5	AAU79864 Human cys
7	273	42.7	48	5	AAU79867 Human cys
8	272	42.5	49	5	AAU79863 Human cys
9	262	40.9	50	4	AAU15096 Peptide #
10	262	40.9	50	4	ABB34086 Peptide #
11	262	40.9	50	4	AAU27545 Peptide #
12	262	40.9	50	4	ABB32389 Peptide #
13	262	40.9	50	4	ABB28913 Peptide #
14	262	40.9	50	4	ABB19524 Protein #
15	262	40.9	50	4	AAU67252 Human bon
16	262	40.9	50	4	AAU54871 Human bra
17	262	40.9	50	4	ABGA48915 Human liv
18	262	40.9	50	4	AAU02833 Peptide #
19	262	40.9	50	5	ABG36903 Human pep
20	254	39.7	46	5	AAU79860 Human cys
21	214	33.4	142	4	AAE02404 Murine cy
22	214	33.4	142	4	AAE04433 Mouse cys
23	214	33.4	143	6	ADA14374 Mouse spe
24	199	31.1	142	7	ADD46708 Rat Prote
25	199	31.1	142	7	ADD46704 Rat Prote

ALIGNMENTS

RESULT 1
AAU79854
ID AAU79854 standard; protein; 117 AA.

XX AAU79854;

XX AC

XX DT 15-JUL-2002 (first entry)

XX DE Human cystatin-8 (zcys8) antigenic fragment #2.

XX KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

XX KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

XX KW sperm motility; fertilisation; antigenic fragment.

XX OS Homo sapiens.

XX PN WO200220567-A2.

XX XX 14-MAR-2002.

XX PF 29-AUG-2001; 2001WO-US026868.

XX PR 01-SEP-2000; 2000US-0230230P.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Holloway JL, Gao Z, Bishop PD;

XX DR WPI; 2002-383044/41.

XX PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

XX PS Claim 2; Page 94-95; 100pp; English.

XX CC The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large

AAU79862 Human cys
AAU79856 Human cys
AAU78260 Fragment
AAU96576 Murine cy
AAU04403 Murine cy
AAU04432 Mouse tes
AAU78258 Fragment
AAU02405 Human cys
AAU04434 Human cys
AAU57231 Human sec
AAU41112 Human sec
AAU74335 Human sec
AAU37980 Human sec
AAU46706 Human Pro
AAU46710 Human Pro
AAU79858 Human cys
AAU78259 Fragment
AAU04315 Alternati
AAU76555 Human Zcy
ABG75917 Human cys

33 5 AAU79862
36 5 AAU79856
123 2 AAU78260
141 3 AAU96576
141 4 AAU02403
141 4 AAU04432
142 2 AAU78258
142 4 AAU02405
142 4 AAU04434
142 6 AAU57231
142 6 AAU41112
142 7 AAU74335
142 7 AAU37980
142 7 AAU46706
142 7 AAU46710
35 5 AAU79858
92 2 AAU78259
145 4 AAU04315
145 5 AAU76555
145 6 ABG75917

26 189 29.5
27 189 29.5
28 188.5 29.5
29 188.5 29.5
30 188.5 29.5
31 188.5 29.5
32 188.5 29.5
33 188.5 29.5
34 188.5 29.5
35 188.5 29.5
36 188.5 29.5
37 188.5 29.5
38 188.5 29.5
39 188.5 29.5
40 188.5 29.5
41 187 29.2
42 184.5 28.8
43 177.5 27.7
44 177.5 27.7
45 177.5 27.7

Query Match	17.9%	Score 112.5;	DB 11;	Length 146;
Best Local Similarity	26.1%;	Pred. No. 0.00027;		
Matches	23;	Conservative 20;	Mismatches 38;	Indels 7; Gaps 1;
Qy	32	YNKESDKYHFRIRVLVKVQRQVTDHLHYHLNVEMQWTTCKP-----ETTNCVPOER	84	
Db	56	YNMGSDSLIYPRDTKVDAKYLQVLVAGIKYYLTLDISTECSRTRVSGEHMDLTTCTLAAG	115	
Qy	85	ELHQVNCFVSFPAVPHFEEQYKILNKSC	112	
Db	116	GQQBKLRGNFELLSEVPWKNTTQLAKHDC	143	

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).
 DR EMBL; AK020193; BAB32024.1; --
 DR HSSP; P01034; IG96.
 DR MGD; MGI:1925859; 8030411F24Rik.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 130 AA; 14947 MW; DD2F930B64B4E584 CRC64;

Query Match 19.2%; Score 121; DB 11; Length 130;
 Best Local Similarity 27.5%; Pred. No. 2.9e-05;
 Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 6 KTFLSVHEVMVENYAKDSLOWITDQYNKESDDKYHFRIVLKVQVQVTDHLYHLNVE 65

DB 23 KEFLDVTKDL--DYFVASVEFAVAQFNDNNPEENTYKLLVGRQKK-TWTMIFLMDLE 78

QY 66 MQWTTCCQKETT--NCVPQERLHKOVNCFSSFAVPWFQYKILNKSC 112

DB 79 MGRITCKHDENIHNCPLQSGREKKVHCVFQVDARPFWSHFILTSTC 127

RESULT 12

Q8VII3 ID Q8VII3 PRELIMINARY; PRT; 130 AA.

AC Q8VII3

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Cystatin SC.

GN 8030411F24Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family";

RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF440735; AAL30841.1; --

DR MGD; MGI:1925859; 8030411F24Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; CY; 1.

SQ SEQUENCE 130 AA; 15076 MW; DD34930B64AF5E8F CRC64;

Query Match 19.2%; Score 121; DB 11; Length 130;

Best Local Similarity 27.5%; Pred. No. 2.9e-05;

Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY 6 KTFLSVHEVMVENYAKDSLOWITDQYNKESDDKYHFRIVLKVQVQVTDHLYHLNVE 65

DB 23 KEFLDVTKDL--DYFVASVEFAVAQFNDNNPEENTYKLLVGRQKK-TWTMIFLMDLE 78

QY 66 MQWTTCCQKETT--NCVPQERLHKOVNCFSSFAVPWFQYKILNKSC 112

DB 79 MGRITCKHDENIHNCPLQSGREKKVHCVFQVDARPFWSHFILTSTC 127

RESULT 13

Q9DAN8 ID Q9DAN8 PRELIMINARY; PRT; 128 AA.

AC Q9DAN8

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE 1700006F03Rik protein (Cystatin TE-1).

GN 1700006F03Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsumoto Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57; TISSUE=Testis;

RA Li Y., Friel P.J., Griswold M.D.;

RT "Molecular cloning and characterization of cystatin SC and cystatin

TE-1, new members of the cystatin family";

RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK005670; BAB24179.1; --

DR EMBL; AF440737; AAL30843.1; --

DR MGD; MGI:1916612; 1700006F03Rik.

DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; CY; 1.

SQ SEQUENCE 128 AA; 15036 MW; 40AFD00103E5E5D6 CRC64;

Query Match 18.2%; Score 114.5; DB 11; Length 128;

Best Local Similarity 29.7%; Pred. No. 0.00014;

Matches 27; Conservative 20; Mismatches 41; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIVLKVQVQVTDHLYHLNVEQWTTCCQK--PETTNCVP 81

DB 36 SVEHVVFHFNENQDDDFAYKFLRVRSLRQKYT-LKYLVDLEMGRTLCGKYDEDIDNCPL 94

QY 82 QERELHKOVNCFSSFAVPWFQYKILNKSC 112

DB 95 QEGPGERKVRCTYIVETEAWTKFILNSTC 125

RESULT 14

Q9JMB4 ID Q9JMB4 PRELIMINARY; PRT; 148 AA.

AC Q9JMB4

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE DD72 protein (Similar to cystatin 10) (Chondrocytes).

GN CST10 OR DD72

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.


```

RN  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  Bai J., Lao H., Ye X., Li Y., Lou J.;
RT  "Molecular cloning and sequence analysis of cystatin cDNA from two
RL  species of sturgeons.";
DR  EMBL; AF334610; AAK16731.1; -.
DR  HSSP; P01038; 1A90.
DR  GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR  InterPro; IPR000010; Cystatin.
DR  Pfam; PF00031; cystatin; 1.
DR  SMART; SM00043; CY; 1.
DR  PROSITE; PS00287; CYSTATIN; 1.
FT  NON TER
SQ  SEQUENCE 112 AA; 12231 MW; 48CEBFBED8A08C00 CRC64;

Query Match      22.1%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 2.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY  18 ENYKDSLOWITDQYNKESDDKHFRIFRVLKQVQVTDHLEHVLNVEQWTTTCQK---P 74
DB  13 EGVQDALKEFAVEAFKASNDMTIHRVSKVKVQKQVAGIKYIVTQMGRTSCRGGAE 72

QY  75 ETTNC-----VPOBELHKQVNCFFSVFVAPVFPFQYKILNKSCS 113
DB  73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLVKNKTC 112

RESULT 9
Q98SR3 ID Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser schrenckii (Amur sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
DR EMBL; AF334611; AAK16732.1; -.
DR HSSP; P01038; 1A90.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFBED8A08C00 CRC64;

Query Match      22.1%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 2.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY  18 ENYKDSLOWITDQYNKESDDKHFRIFRVLKQVQVTDHLEHVLNVEQWTTTCQK---P 74
DB  13 EGVQDALKEFAVEAFKASNDMTIHRVSKVKVQKQVAGIKYIVTQMGRTSCRGGAE 72

QY  75 ETTNC-----VPOBELHKQVNCFFSVFVAPVFPFQYKILNKSCS 113
DB  73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLVKNKTC 112

RESULT 10
Q98SR3 ID Q98SR3 PRELIMINARY; PRT; 112 AA.
AC Q98SR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser schrenckii (Amur sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=111304;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bai J., Lao H., Ye X., Li Y., Lou J.;
RT "Molecular cloning and sequence analysis of cystatin cDNA from two
RL species of sturgeons.";
DR EMBL; AF334611; AAK16732.1; -.
DR HSSP; P01038; 1A90.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON TER
SQ SEQUENCE 112 AA; 12231 MW; 48CEBFBED8A08C00 CRC64;

Query Match      22.1%; Score 139; DB 13; Length 112;
Best Local Similarity 32.0%; Pred. No. 2.9e-07;
Matches 33; Conservative 24; Mismatches 36; Indels 10; Gaps 3;

QY  18 ENYKDSLOWITDQYNKESDDKHFRIFRVLKQVQVTDHLEHVLNVEQWTTTCQK---P 74
DB  13 EGVQDALKEFAVEAFKASNDMTIHRVSKVKVQKQVAGIKYIVTQMGRTSCRGGAE 72

QY  75 ETTNC-----VPOBELHKQVNCFFSVFVAPVFPFQYKILNKSCS 113
DB  73 KIELCAFDHPV---ELAKTSTCTFEVVSRLWIPETKLVKNKTC 112
```

```

Q8VIH8 ID Q8VIH8 PRELIMINARY; PRT; 130 AA.
AC Q8VIH8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin SC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RA Li Y., Friel P.J., Griswold M.D.;
RT "Molecular cloning and characterization of cystatin SC and cystatin
RL TE-1, new members of the cystatin family.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF442205; AAL35350.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR SEQUENCE 130 AA; 14981 MW; 7A752359860989C9 CRC64;

Query Match      19.4%; Score 122; DB 11; Length 130;
Best Local Similarity 27.5%; Pred. No. 2.3e-05;
Matches 30; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

QY  6 KTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKHFRIFRVLKQVQVTDHLEHVLNVE 65
DB  23 KEFLDVTQDQ---DYFVASVEFAVAFQNDNNSEENTYRLLEVGRAQKK-TWTMIFLMDLE 78

QY  66 MQWTTCCQKPETT--NCVPOERELHKQVNCFFSVFVAPVFPFQYKILNKSC 112
DB  79 MERTICKKHDEHNHCPLLOGSGEKKVCHVCVQVDARPFWSHFTVLSTC 127

RESULT 11
Q9CX46 ID Q9CX46 PRELIMINARY; PRT; 130 AA.
AC Q9CX46;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 8030411F24Rik protein.
GN 8030411F24Rik
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/60; TISSUE=Embryonic testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Niehi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Kasavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
```

```
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AA48681.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A894951F CRC64;

Query Match      29.8%; Score 187.5; DB 11; Length 141;
Best Local Similarity 39.6%; Pred. No. 2.3e-12;
Matches 38; Conservative 20; Mismatches 35; Indels 3; Gaps 2;

QY 20 YAKDSLOWITDQYNKESDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC--PETT 77
Db 43 YVOHALWYAMKEYNKAASNDLYNFRVVDILKSQSDITDSLEYLYLEVNARTMCKKIAGDNE 102

QY 78 NCV-POBERELHKQVNCFFSVFVFPWPFQYKILNKSC 112
Db 103 NCLFQDDPKMKKXVFCIFIVSSKPFKELKWLKQC 138

RESULT 6
Q9EPX9 PRELIMINARY; PRT; 140 AA.
AC Q9EPX9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=21010502; PubMed=11144350;
RA Taupin P.J., Ray J., Fischer W.H., Suhr S.T., Hakansson K., Grubb A.,
RA Gage F.H.;
RT "FGF-2-Responsive neural stem cell proliferation requires Ccg, a novel
RL autocrine/paracrine cofactor.";
RL Neuron 28:385-397(2000).
DR EMBL; AF311741; AAC40283.1; -.
DR HSP; P01034; I996.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT CHAIN 21 140 CYSTATIN C.
FT VARIANT 16 16 A -> G.
FT VARIANT 84 84 L -> F.
SQ SEQUENCE 140 AA; 15517 MW; 3A563406DD58D785 CRC64;

Query Match      26.5%; Score 166.5; DB 11; Length 140;
Best Local Similarity 33.3%; Pred. No. 4.1e-10;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

QY 13 EVMAVENYAKDSLOWITDQYNKESDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC 72
Db 35 EADANEGVRALDFAVSEYKNGSNDAYHSRAQVVRARQQLVAGVNVFLDVGRTTCT 94

QY 73 KPET--TNC-VPOBERELHKQVNCFFSVFVFPWPFQYKILNKSCSS 114
Db 95 KSOQNLTDCEFDHQPMLRRALCSFOIYSVPWKGTSLTNFSCN 139

RESULT 7
Q80Y72 PRELIMINARY; PRT; 140 AA.
ID Q80Y72
Q80Y72
```

```
AC Q80Y72;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cystatin-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22388237; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Scheetz T.E.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Prange C.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullaby S.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Munaratne P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048646; AA48646.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 140 AA; 16199 MW; 32633E99C4697DA0 CRC64;

Query Match      23.8%; Score 149.5; DB 11; Length 140;
Best Local Similarity 30.4%; Pred. No. 2.7e-08;
Matches 31; Conservative 27; Mismatches 41; Indels 3; Gaps 2;

QY 13 EVMAVENYAKDSLOWITDQYNKESDDKYHFRIFVLKQVQVTDHLEYHLNVEMQWTTCC 72
Db 37 KAMSKKN-INSTLHFFIRSYNNASNDYLYQVKLIQGGQMLTGVLYLVTVIGRTCK 95

QY 73 KPET--TNCVPOBERELHKQVNCFFSVFVFPWPFQYKILNKSC 112
Db 96 KNETKASCPOSSKLUKSLICKSLIYVPMWNYQLWNNSC 137

RESULT 8
Q98SR4 PRELIMINARY; PRT; 112 AA.
ID Q98SR4
AC Q98SR4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin (Fragment).
OS Acipenser sinensis (Chinese sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=61970;
```


Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	629	100.0	138	4	Q8WXU6	Q8wxu6 homo sapien
2	400.5	63.7	103	4	Q8WXU5	Q8wxu5 homo sapien
3	360	57.2	139	11	Q8KS53	Q8ks53 rattus norv
4	187.5	29.8	141	11	Q8DAP1	Q8dap1 mus musculus
5	187.5	29.8	141	11	Q8OZM5	Q8ozm5 mus musculus
6	166.5	26.5	140	11	Q8EPX9	Q8epx9 mus musculus
7	149.5	23.8	140	11	Q8OY72	Q8oy72 mus musculus
8	139	22.1	112	13	Q8ESR4	Q8esr4 acipenser s
9	139	22.1	112	13	Q8ESR3	Q8esr3 acipenser s
10	122	19.4	130	11	Q8VIH8	Q8vih8 rattus norv
11	121	19.2	130	11	Q8CX46	Q8cx46 mus musculus
12	121	19.2	130	11	Q8VII3	Q8vii3 mus musculus
13	114.5	18.2	128	11	Q8DAN8	Q8dan8 mus musculus
14	113	18.0	148	11	Q9JMR4	Q9jmr4 mus musculus
15	112.5	17.9	146	11	Q8K397	Q8k397 mus musculus
16	112.5	17.9	149	11	Q8VHC1	Q8vhc1 rattus norv


```

AC O19092;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CST3.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
RL cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U51912; BAB64050.1; -.
DR HSP; P01034; I396.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29DF26 CRC64;

Query Match 24.6%; Score 155; DB 1; Length 146;
Best Local Similarity 33.7%; Pred. No. 5.1e-09;
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;

QY 18 ENYAKSLQWITDQYNKESDDKTHFRIVLVKVRQVTDHLEYHLNVEMQWTTQK--PE 75
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
46 EGVRRALDPVAVSEYKNSNDYHSRALQVVRARQIVAGVYFLDVELGRITCTKTPN 105
QY 76 TTNCVQER-ELHKQVNCFFSVAPWPFQYKILNKSCSSD 115
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
106 LDNCPFHEQPHLRKKAFCFQIVTPW-QGTMTLSKSTCD 145

```

Search completed: March 18, 2004, 14:16:08
Job time : 14.6905 secs

RT "Characterization and amino acid sequence of a new acidic cysteine
RT proteinase inhibitor (cystatin SA) structurally closely related to
RT cystatin S, from human whole saliva.";
RL J. Biochem. 102:693-704(1987).
RN [5]
RP PRELIMINARY SEQUENCE OF 25-141.
RA Isemura S., Saitoh E., Sanada K., Isemura M., Ito S.;
RT "Characterization and amino acid sequence of a new acidic cysteine
RT proteinase inhibitor (cystatin SA) structurally closely related to
RT cystatin S, from human whole saliva.";
RL (in Turk V. (eds.);
RL Cysteine proteinases and their inhibitors, pp.497-505,
RL Walter de Gruyter, Berlin and New York (1986).
RN [6]
RP SEQUENCE OF 25-141 FROM N.A.
RX MEDLINE=89076505; PubMed=3202964;
RA Saitoh E., Isemura S., Sanada K., Kim H.-S., Smithies O., Maeda N.;
RT "Cystatin superfamily. Evidence that family II cystatin genes are
RT evolutionarily related to family III cystatin genes.";
RL Biol. Chem. Hoppe-Seyler 369:191-197(1988).
CC -!- FUNCTION: Thiol protease inhibitor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M19673; AAA36116.1; -.
DR EMBL; M19671; AAA36116.1; JOINED.
DR EMBL; M19672; AAA36116.1; JOINED.
DR EMBL; AL591074; CAC94784.1; -.
DR PIR; B29632; B29632.
DR HSSP; P01034; 1G96.
DR Genew; HGNC:2474; CST2.
DR MIM; 123856; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal; Multigene family.
FT SIGNAL 1 20
FT CHAIN 21 141 CYSTATIN SA.
FT ACT_SITE 32 32 REACTIVE SITE.
FT SITE 76 80 SECONDARY AREA OF CONTACT.
FT DISULFID 94 104 BY SIMILARITY.
FT DISULFID 118 138 BY SIMILARITY.
SQ SEQUENCE 141 AA; EB54915B1B977AA2 CRC64;
Query Match 25.5%; Score 160.5; DB 1; Length 141;
Best Local Similarity 31.5%; Pred. No. 1.3e-09;
Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;
QY 24 SLQWITDQYNKESDDKYHFRIFRLVKVQVOTDHLVHLNVMQWTTCK--PETTNCVP 81
DB 47 ALHFVISEYNKATEDEYRLLRLRAREQIVGVNFFDIEVGRITCTKSNLDTCF 106
QY 82 QER-ELHKQVNCFFSFVAFVFWFEQYKILNKSC 112
DB 107 HEQPELQKQLCSFQIYVFWDEMSLVNSRC 138
RESULT 12
CYTC SAISC STANDARD; PRT; 146 AA.
AC O19093;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CST3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U52028; AAB64051.1; -.
DR HSSP; P01034; 1G96.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 08196353C0306AA3 CRC64;
Query Match 25.4%; Score 160; DB 1; Length 146;
Best Local Similarity 34.7%; Pred. No. 1.6e-09;
Matches 35; Conservative 21; Mismatches 41; Indels 4; Gaps 3;
QY 18 ENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQVOTDHLVHLNVMQWTTCK--PE 75
DB 46 EGVRRALDFAVSEYNKASNDMYSRALQVVRARKQIVAGVYFLDVEMGRITCTKNQPN 105
QY 76 TTNCVPQER-ELHKQVNCFFSFVAFVFWFEQYKILNKSCSSD 115
DB 106 LDNCPFHEPFLKRAKAFCSFQIYVDPV--QGIMTLKSKTCQD 145
RESULT 13
CSTL HUMAN STANDARD; PRT; 165 AA.
ID _CSTL_HUMAN
AC QSH114;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin-like 1 precursor.
GN CSTL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coughlin A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,
 RA Grahman D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.B., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leheslahti M.H., Leversha M.A., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,
 RA Swann R.M., Sycamore A.C., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation. LOCATION: Secreted.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF059244; AAC14707.1; -;
 CC ENBL; AL109954; CAB64234.1; -;
 CC HSSP; P01034; IG96.
 CC Genew; HGNC:2480; CST8.
 CC GO; GO:0004869; F:Cysteine protease inhibitor activity; TAS.
 CC InterPro; IPR000010; Cystatin.
 CC Pfam; PF00031; cystatin; 1.
 CC SMART; SM00043; CY; 1.
 CC Thiol protease inhibitor; Signal; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 142
 FT SITE 77 81
 FT SITE 95 105
 FT DISULFID 119 139
 FT DISULFID 27 27
 FT CARBOHYD 39 39
 FT VARIANT 142 142
 FT SEQUENCE 142 AA; 16275 MW; 9A3512757E0F4ECD CRC64;
 Query Match 30.0%; Score 188.5; DB 1; Length 142;
 Best Local Similarity 40.4%; Pred. No. 1.8e-12;
 Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;
 QY 22 KDSLOWITDOYNKESDDKYHFRVLKQVROVTDHLEHVLNVMQWTTCKPETTNC-C 79
 DB 46 KQCLWFAMQYNKESDDKYHFRVLKQVROVTDHLEHVLNVMQWTTCKPETTNC 105
 QY 80 VPQER-ELHKQVNCFFSVFVFPVFEQYKILNKC 112
 DB 106 AIQENSKLKKLSCFVLGALPWNGETVMEKCC 139

RESULT 6
 CYTC_RABIT STANDARD; PRT; 148 AA.
 ID CYTC_RABIT
 AC O97862;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor.
 GN CST3.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCBI_TaxID=9986;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Japanese white; TISSUE=Bone;
 RX MEDLINE=98424349; PubMed=9753427;
 RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,
 RA Kawashima H.;
 RT "Large scale isolation of osteoclast-specific genes by an improved
 RT method involving the preparation of a subtracted cDNA library.";
 RL Genes Cells 3:459-475(1998).
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB009342; BAA75921.1; -;
 CC HSSP; P01034; IG96.
 CC InterPro; IPR000010; Cystatin.
 CC Pfam; PF00031; cystatin; 1.
 CC SMART; SM00043; CY; 1.
 CC PROSITE; PS00287; CYPSTATIN; FALSE_NEG.
 CC Thiol protease inhibitor; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 148
 FT ACT_SITE 39 39
 FT SITE 83 87
 FT DISULFID 101 111
 FT DISULFID 125 145
 FT SEQUENCE 148 AA; 16346 MW; 1523C8311695B9A CRC64;
 Query Match 27.7%; Score 174.5; DB 1; Length 148;
 Best Local Similarity 30.8%; Pred. No. 5.2e-11;
 Matches 33; Conservative 30; Mismatches 41; Indels 3; Gaps 2;
 QY 11 VHEVMAVENYAKDSLOWITDOYNKESDDKYHFRVLKQVROVTDHLEHVLNVMQWTT 70
 DB 41 LBDVDAQKDVQKALGFAEYKNGSNDYHRSALQVVRARRQIVSGVKYLDVIGRTT 100
 QY 71 CQKPEP--TNC-VPOREHLKQVNCFFSVFVFPVFEQYKILNKCSCS 114
 DB 101 CTKTQTNLANCPHDPDQLQKMLCSFEIYVPLNKLISLLKSDCCN 147
 RESULT 7
 CYTC_RAT STANDARD; PRT; 127 AA.
 ID CYTC_RAT
 AC P14841;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor (Fragment).
 GN CST3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```
CC maturation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
CC expression in the testis. Within the testis it is localized to the
CC elongating spermatids, whereas within the epididymis it is
CC exclusively synthesized by the proximal caput epithelium.
CC -!- INDUCTION: Testicular factors or hormones other than androgens
CC present in the testicular fluid may be involved in the regulation
CC of CRE gene expression.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF091503; AAC61754.1; -.
CC EMBL; AF090691; AAC36316.1; -.
CC EMBL; S49926; AAC35390.1; -.
CC FIRM; A45361; A45361.
CC HSSP; P01034; IG96.
CC MGD; MGI:107161; Cst8.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; cystatin; 1.
CC SMART; SM00043; Cy; 1.
CC Thiol protease inhibitor; Signal.
CC SIGNAL 1 19
CC CHAIN 20 142
CC SITE 77 81 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
CC DISULFID 95 105 PROTEIN.
CC DISULFID 119 139 SECONDARY AREA OF CONTACT (POTENTIAL).
CC CARBOHYD 39 39 BY SIMILARITY.
CC CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CONFLICT 4 15 PLWSLILFTIP -> GTRDEOVGESQK (IN REF. 2).
CC SEQUENCE 142 AA; 16288 MW; 50B446E98F66673E CRC64;
CC
CC Query Match 34.0%; Score 214; DB 1; Length 142;
CC Best Local Similarity 37.1%; Pred. No. 4.5e-15;
CC Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;
CC
CC QY 10 SVHEWAVENY-----AKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLE 59
CC DB 25 SKNEVKA-QNYFSINISNANVKQVWFMKYEKSEDKYFLVDKILHAKLQITDRME 83
CC
CC QY 60 YHLNVEMQWTTCKP--ETTNCVPOER-ELHKQVNCFFSVFVFPWFQYKILNKC 112
CC DB 84 YQIDVQISRNCCKPLNNTENCIPQKPKPELEKXKSCFLVGLPWNCFNLLSKEC 139
CC
CC RESULT 4
CC CSTR8_RAT
CC ID CSTR8_RAT STANDARD; PRT; 142 AA.
CC AC O88969;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
CC DE 8).
CC GN CSTR8 OR CRES.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX STRAIN=Sprague-Dawley; TISSUE=Epididymis;
CC RX MEDLINE=99247899; PubMed=10229662;
CC RA Cornwall G.A., Hsiao N., Sutton H.G.;
CC RA "Structure, alternative splicing and chromosomal localization of the
```

```
RT cystatin-related epididymal spermatogenic gene.";
RL Biochem. J. 340:85-93(1999).
CC -!- FUNCTION: Performs a specialized role during sperm development and
CC maturation.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF090692; AAC36317.1; -.
CC HSSP; P01034; IG96.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; cystatin; 1.
CC SMART; SM00043; Cy; 1.
CC Thiol protease inhibitor; Signal.
CC SIGNAL 1 19
CC CHAIN 20 142
CC SITE 77 81 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
CC DISULFID 95 105 PROTEIN.
CC DISULFID 119 139 SECONDARY AREA OF CONTACT (POTENTIAL).
CC CARBOHYD 100 100 BY SIMILARITY.
CC SEQUENCE 142 AA; 16246 MW; FB873FAA6B6CAE34 CRC64;
CC
CC Query Match 31.6%; Score 199; DB 1; Length 142;
CC Best Local Similarity 35.7%; Pred. No. 1.5e-13;
CC Matches 41; Conservative 26; Mismatches 36; Indels 12; Gaps 4;
CC
CC QY 10 SVHEWAVENYAKDSLQ-----WIT-DQYNKESDDKYHFRIFRVLKQVQVTDHLEY 60
CC DB 25 SKNEVKAQRYFGSISISNANVKQVWFMKYEKSEDKYFLDLTKLTHATLQITDRMEY 84
CC
CC QY 61 HLNVEWQWTTCKP--ETTNCVPOER-ELHKQVNCFFSVFVFPWFQYKILNKC 112
CC DB 85 HLDVQISRNCCKPLNNTENCIPQKPKPELEKXKSCFLVGLPWNCFNLLSKEC 139
CC
CC RESULT 5
CC CSTR8_HUMAN
CC ID CSTR8_HUMAN STANDARD; PRT; 142 AA.
CC AC O60676;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
CC DE 8).
CC GN CSTR8 OR CRES.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Testis;
CC RX MEDLINE=95344753; PubMed=7619504;
CC RA Cornwall G.A., Hann S.R.;
CC RT "Transient appearance of CRES protein during spermatogenesis and
CC RL caput epididymal sperm maturation.";
CC RL Mol. Reprod. Dev. 41:37-46(1995).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=21638749; PubMed=11780052;
CC RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
CC RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
CC RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
CC RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
CC RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AL096677; CAC13170.1; -.
 HSSP; AL096677; CAC17423.1; -.
 HSSP; P01038; IAG0.
 Genew; HGNC:15959; CSTR11.
 InterPro; IPR000010; Cystatin.
 Pfam; PF00031; cystatin; 1.
 SMART; SM00043; Cy; 1.
 PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 Thiol protease inhibitor; Signal; Alternative splicing.
 SIGNAL 1 25 POTENTIAL.
 CHAIN 26 137 SECONDARY AREA OF CONTACT (POTENTIAL).
 SITE 75 79 BY SIMILARITY.
 DISULFID 93 101 BY SIMILARITY.
 DISULFID 114 134 BY SIMILARITY.
 CARBOHYD 131 131 N-LINKED (GLCNAC...) (POTENTIAL).
 VARSPLIC 76 110 Missing (in isoform 2).
 FTID=VSP 001260.
 SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 629; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.5e-57;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFSLVHEVMAVENYAKDSQWITDQYNKESDDKYHFRFVLKQVQVOTDHELY 60
 DB 23 YQARKKTFSLVHEVMAVENYAKDSQWITDQYNKESDDKYHFRFVLKQVQVOTDHELY 82
 QY 61 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCSSD 115
 DB 83 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCSSD 137

RESULT 2

CS11_MOUSE STANDARD; PRT; 139 AA.
 AC Q9D269;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin 11 precursor.
 GN CSTR11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Epididymis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AK020300; BAB32061.1; -.
 HSSP; P01034; IG96.
 MGD; MGI:1925490; Cst11.
 InterPro; IPR000010; Cystatin.
 Pfam; PF00031; cystatin; 1.
 SMART; SM00043; Cy; 1.
 PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 Thiol protease inhibitor; Signal.
 SIGNAL 1 28 POTENTIAL.
 CHAIN 29 139 SECONDARY AREA OF CONTACT (POTENTIAL).
 SITE 76 80 BY SIMILARITY.
 DISULFID 94 102 BY SIMILARITY.
 DISULFID 115 135 BY SIMILARITY.
 CARBOHYD 134 134 N-LINKED (GLCNAC...) (POTENTIAL).
 SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;

Query Match 59.1%; Score 372; DB 1; Length 139;
 Best Local Similarity 54.9%; Pred. No. 3e-31;
 Matches 62; Conservative 29; Mismatches 22; Indels 0; Gaps 0;

QY 1 YQARKKTFSLVHEVMAVENYAKDSQWITDQYNKESDDKYHFRFVLKQVQVOTDHELY 60
 DB 24 YQARKKTFIRIEVSALESSVKTLEYVTDEYNKESDLYNFRILKIMKQVTHLEY 83
 QY 61 HLNVENQWTTCKPPTTNCVQPERLHKQVNCFFSVFVFPWFQYKILNKSCS 113
 DB 84 HITVEMQRTTCLKTSTSLCDIQKGLHKKIQCYFSVYAIIPWVEVFKILKNCT 136

RESULT 3

CS18_MOUSE STANDARD; PRT; 142 AA.
 AC P32766; O89102;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-
 DE related epididymal specific protein) (Cystatin 8).
 GN CST8 OR CRES.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C3H, and CD-1;
 RX MEDLINE=99247899; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.;
 RT "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermatogenic gene."
 RL Biochem. J. 340:85-93(1999).
 RN [2]
 RP SEQUENCE OF 4-142 FROM N.A.
 RC TISSUE=Epididymis;
 RX MEDLINE=93078799; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;
 RT "The CRES gene: a unique testis-regulated gene related to the cystatin
 RT family is highly restricted in its expression to the proximal region
 RT of the mouse epididymis."
 RL Mol. Endocrinol. 6:1653-1664(1992).
 CC -!- FUNCTION: Performs a specialized role during sperm development and

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 13.6905 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-3
Perfect score: 629
Sequence: 1 YQARKKTFSLVHEVMAVENVY.....VFAVPWFQYKLNKSCSD 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	629	100.0	137	CS11 HUMAN	Q9h112 homo sapien
2	372	59.1	139	CS11 MOUSE	Q9d269 mus musculus
3	214	34.0	142	CS18 MOUSE	P32766 mus musculus
4	199	31.6	142	CS18 RAT	O88969 rattus norv
5	188.5	30.0	142	CS18 HUMAN	O60676 homo sapien
6	174.5	27.7	148	CS18 RABBIT	O97862 oryctolagus
7	171.5	27.3	127	CYTC RAT	P14841 rattus norv
8	170.5	27.1	139	CYTC CHICK	P01038 gallus gall
9	169.5	26.9	116	CYT CORJA	P81061 coturnix co
10	160.5	25.5	140	CYTC MOUSE	P21460 mus musculus
11	160.5	25.5	141	CYTC HUMAN	P09228 homo sapien
12	160	25.4	146	CYTC SAISC	O19093 salmisi sci
13	157.5	25.0	165	CS18 HUMAN	Q9h114 homo sapien
14	155	24.6	146	CS18 HUMAN	P01034 homo sapien
15	155	24.6	146	CYTC MACMU	O19092 macaca mula
16	153.5	24.4	148	CYTC BOVIN	P01035 bos taurus
17	146.5	23.3	141	CYTC HUMAN	P01036 homo sapien
18	143	22.7	129	CYT CYPCA	P35481 cyprinus ca
19	139.5	22.2	141	CYTN HUMAN	P01037 homo sapien
20	135	21.5	147	CS19 HUMAN	Q9h491 homo sapien
21	130.5	20.7	142	CYTD HUMAN	P28325 homo sapien
22	129.5	20.6	141	CYTS RAT	P19313 rattus norv
23	127	20.2	111	CYT BITAR	P08935 bitis ariet
24	127	20.2	130	CYT ONCMY	O91195 oncorhynch
25	127	20.2	137	CS19 MOUSE	Q9z0h6 mus musculus
26	126	20.0	130	CYT ONCKE	Q98967 oncorhynch
27	121.5	19.3	149	CYTM HUMAN	O15828 homo sapien
28	109	17.3	145	CYTF HUMAN	O76096 homo sapien
29	98.5	15.7	434	1 KNL2 BOVIN	P01047 bos taurus
30	98.5	15.7	619	1 KNL2 BOVIN	P01045 bos taurus
31	98	15.6	144	1 CYTF MOUSE	O89098 mus musculus
32	96.5	15.3	436	1 KNL1 BOVIN	P01046 bos taurus
33	96.5	15.3	621	1 KNL1 BOVIN	P01044 bos taurus

34	95.5	15.2	162	1	CYTX ONCVO	P22085 onchocerca
35	93.5	14.9	644	1	KNG HUMAN	P01042 homo sapien
36	87	13.8	639	1	KNG RAT	P08934 rattus norv
37	82.5	13.1	474	1	SYE_VIBPA	O87r16 vibrio para
38	81	12.9	430	1	KNT2 RAT	P08932 rattus norv
39	79	12.6	430	1	KNT1 RAT	P01048 rattus norv
40	79	12.6	661	1	KNG MOUSE	O08677 mus musculus
41	77	12.2	214	1	CAPA_KLUULA	O74232 kluyveromyc
42	77	12.2	430	1	MESJ_HAEIN	P44689 haemophilus
43	75.5	12.0	474	1	SYE_VIBVY	Q7mmw8 vibrio vuln
44	75	11.9	400	1	BGIB HUMAN	Q06430 homo sapien
45	73.5	11.7	438	1	EXG_CANAL	P29717 candida alb

ALIGNMENTS

```
RESULT 1
CS11_HUMAN STANDARD; PRT; 137 AA.
AC Q9H112; Q9H113;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Cystatin 11 precursor.
GN CST11 OR CST8L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knighs A., Laird G.K., Lawlor S.,
RA Lehaeslaio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9H112-1; SequencesDisplayed;
CC Name=2;
CC IsoId=Q9H112-2; SequencesVSP 001260;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


Query Match	20.2%;	Score 127;	DB 2;	Length 111;
Best Local Similarity	30.3%;	Pred. No. 1.1e-05;		
Matches	27;	Conservative 24;	Mismatches 28;	Indels 10; Gaps 3;
Qy	22	KDSLQWITDQYNKESDDKYHFRFVLVKQQRQVTDHLYHLNVEMQWTTQK-----	73	
Db	16	QEAFAVEKYNAGSKNDYYFKERRVVEAQSVGVKYIIMBELLTKTKTGRPKGY	75	
Qy	74	PETTNC-VPOERELHKQVNCFFSVFAVPW	101	
Db	76	QEIQCNLPPENQ-QEITCRFEVMSRPW	103	

RESULT 15

Cystatin precursor - chum salmon
 C:Species: Oncorhynchus keta (Chum salmon)
 C:date: 28-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jan-2000
 C:Accession: JC4918
 R:Yamashita, M.; Konagaya, S.
 J. Biochem. 120, 483-487, 1996.
 A:title: Molecular cloning and gene expression of chum salmon cystatin.
 A:Reference number: JC4918; PMID:8902609
 A:Accession: JC4918
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-132 <YAM>
 A:Cross-references: DDBJ:D86628
 A:Experimental source: liver
 C:Comment: This protein is a homolog of mammalian cystatin C, and acts as an extracellular
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: liver
 F:/1-21/Domain: signal sequence #status predicted <SIG>
 F:/23-132/Domain: cystatin homology <CYS>

	Query Match	20.08;	Score 126;	DB 2:	Length 132;	
	Best Local Similarity	26.3%;	Pred.No. 1.6e-05;			
	Matches	25;	Conservative	41;	Indels	2; Gaps 1;
Qy	22 KDSLOWITDOYNKESDDKHFRFRLUKVQRQVTDHLEYHLNVEVMQWTCOK--PETTNC	79				
Dd	38 RDLAQFAVWEKKNTNDMFVRQAKVVNAQAQQVVSGMKYIFTVMGRTPCRKGGVEXICS	97				
Qy	80 VPQERLHKQVCNCFFSFVAPVPWFPEQYKILNKSCSS	114				
Dd	98 VHKDPQMAVPYKCTFEVMSRPNSDIOMNKNOCES	132				

Search completed: March 18, 2004, 14:23:12
Job time : 23.1329 secs

A;Cross-references: GB:J03870; NID:g337751; PIDN:AAA60299.1; PID:g337752
R;Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988
A;Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily
A;Reference number: S02489; MUID:89076505; PMID:3202964
A;Accession: S02489
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 21-141 <SA2>
R;Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.
Gene 61, 329-338, 1987
A;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem
A;Reference number: A91589; MUID:89185836; PMID:3446578
A;Accession: A29632
A;Molecule type: DNA
A;Residues: 1-86, 'I', 88-141 <SAI>
R;Isemura, S.; Saitoh, E.; Sanada, K.
FEBS Lett. 198, 145-149, 1986
A;Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystat
A;Reference number: A01273; MUID:86164938; PMID:3514272
A;Accession: A01273
A;Molecule type: protein
A;Residues: 29-141 <ISE>
R;Ramasubbu, N.; Reddy, M.S.; Bergey, E.J.; Haraszthy, G.G.; Soni, S.D.; Levine, M.J.
Biochem. J. 280, 341-352, 1991
A;Title: Large-scale purification and characterization of the major phosphoproteins and
A;Reference number: S19279; MUID:92082469; PMID:1747107
A;Accession: S19279
A;Status: preliminary
A;Molecule type: protein
A;Residues: 21-55 <RAM>
C;Comment: Human saliva appears to contain several cysteine proteinase inhibitors that an
ences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl
C;Genetics:
A;Gene: GDB:CST1
A;Cross-references: GDB:119815; OMIM:123855
A;Map position: 20p11.2-20p11.2
C;Superfamily: cystatin; cystatin homolog
C;Keywords: cysteine proteinase inhibitor; extracellular protein; saliva
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-141/Product: cystatin SA-I #status experimental <MAT1>
F;29-141/Product: cystatin SN #status experimental <MAT2>
F;30-141/Domain: cystatin homolog <CYS>
F;76-80/Region: inhibitory #status predicted
F;94-104,118-138/Disulfide bonds: #status predicted

Query Match 22.2%; Score 139.5; DB 1; Length 141;
Best Local Similarity 29.6%; Pred. No. 7.9e-07;
Matches 29; Conservative 21; Mismatches 45; Indels 3; Gaps 2;
Qy 18 ENYAKDSLQMTTDYKNKESDKYHFRPLVKVQRQVTDHLEYHLNVEMQWTTCKQ--PE 75
Db 41 DEWQRAHFAISYKATKDDYTRRPLRVLRARQQTGGVNYFFDEVEGRICTKSQPN 100
Qy 76 TTNCVQPER-ELHQRVNCFFSVFAPVWPEQYKILNKSC 112
Db 101 LDTCAFHEQPELQKQLCSFEIYEVFWENRRSLVKSRC 138

RESULT 11
JC2040
N;Alternate names: cysteine proteinase inhibitor
C;Species: Oncorhynchus keta (chum salmon)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 16-Jul-1999
C;Accession: JC2040
R;Koide, Y.; Noso, T.
Biosci. Biotechnol. Biochem. 58, 164-169, 1994
A;Title: The complete amino acid sequence of pituitary cystatin from chum salmon.
A;Reference number: JC2040; MUID:94162738; PMID:7764512
A;Accession: JC2040
A;Molecule type: protein
A;Residues: 1-111 <KOI>

A;Accession: A01270
A;Molecule type: protein
A;Residues: 27-131,'S',133-146 <GRU>
R;Ghislo, J.; Jensen, O.; Frangione, B.
Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986
A;Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland
A;Reference number: A25434; MUID:86206076; PMID:3517880
A;Accession: A25434
A;Molecule type: protein
A;Residues: 37-93,'O',95-146 <GHI>
R;Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983
A;Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin
A;Reference number: S01461; MUID:84110059; PMID:6662498
A;Accession: S12288
A;Molecule type: protein
A;Residues: 27-73 <TUR>
R;Brzin, J.; Popovic, T.; Turk, V.
Biochem. Biophys. Res. Commun. 118, 103-109, 1984
A;Title: Human cystatin, a new protein inhibitor of cysteine proteinases.
A;Reference number: A32732; MUID:84128015; PMID:6365094
A;Accession: A32732
A;Molecule type: protein
A;Residues: 27-76 <BRZ>
R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst
A;Reference number: A60552; MUID:90193615; PMID:2315647
A;Accession: A60552
A;Molecule type: protein
A;Residues: 27-49,'XX',52-64 <OLA>
A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
e defective gene is not present in CSF but is found instead in amyloid deposits
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
A;Title: Different forms of human cystatin C.
A;Reference number: S10607; MUID:91025625; PMID:2222856
A;Accession: S10607
A;Molecule type: protein
A;Residues: 27-53 <POP>
A;Experimental source: urine, kidney disease
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als
R;Grubb, A.; Lofberg, H.; Barrett, A.J.
FEBS Lett. 170, 370-374, 1984
A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.
A;Reference number: S01462
A;Contents: annotation; disulfide bonds
R;Berti, P.J.; Storer, A.C.
Biochem. J. 302, 411-416, 1994
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility
A;Reference number: S55305; MUID:94379969; PMID:8092991
A;Accession: S55305
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-49;106-146 <BER>
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl
F patients with certain autoimmune diseases.
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importa
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr
C;Genetics:
A;Gene: GDB:CST3
A;Cross-references: GDB:119817; OMIM:105150
A;Map position: 20p11.2-20p11.2
A;Introns: 81/3; 119/3
C;Superfamily: cystatin; cystatin homology
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyprolin
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-146/Product: cystatin C #status experimental <MAT>
F;35-146/Domain: cystatin homology <CYS>
F;81-85/Region: inhibitory #status predicted
F;29/Modified site: hydroxyproline (Pro) (partial) #status experimental
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 24.6%; Score 155; DB 1; Length 146;
Best Local Similarity 33.7%; Pred. No. 2.3e-08;
Matches 34; Conservative 21; Mismatches 42; Indels 4; Gaps 3;
QY 18 ENYAKDSLOWITDQYNKESDDKHFRIFRVLKQVQRTDHLHYHLNVEMQWTTTCQK--PE 75
DB 46 EGVRRALDPANVEYNKASNDMYHSALQVVRARKQIVAGVNYFLDVLGRITTCCTKQPN 105
QY 76 TTNC-VPOREHLHKQVNCFFSVFAVWFQYKILNKSCSSD 115
DB 106 LDNCPPHDQPHLKRKAFCSQIYAVDW-QGTWTLSTKSTQCD 145
RESULT 8
UDROP1
cystatin - bovine
N;Alternate names: thiol proteinase inhibitor
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 06-Dec-1996
C;Accession: A01271
R;Hirado, M.; Tsunagaawa, S.; Sakiyama, F.; Niinobe, M.; Fujii, S.
FEBS Lett. 186, 41-45, 1985
A;Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase in
A;Reference number: A01271; MUID:85231205; PMID:3891407
A;Accession: A01271
A;Molecule type: protein
A;Residues: 1-112 <HIR>
C;Superfamily: cystatin; cysteine proteinase inhibitor
C;Keywords: colostrum; cysteine proteinase inhibitor
F;2-112/Domain: cystatin homology <CYS>
F;48-52/Region: inhibitory #status predicted
F;66-76,90-110/Disulfide bonds: #status predicted
Query Match 24.4%; Score 153.5; DB 1; Length 112;
Best Local Similarity 30.6%; Pred. No. 2.4e-08;
Matches 30; Conservative 23; Mismatches 42; Indels 3; Gaps 2;
QY 18 ENYAKDSLOWITDQYNKESDDKHFRIFRVLKQVQRTDHLHYHLNVEMQWTTTCQKPE 77
DB 13 EGVQEQALSFVAVSEFNKRNDAVQSRVVRVVRARKQVSGMNYFLDVLGRITTCCTKQAN 72
QY 78 --NC-VPOREHLHKQVNCFFSVFAVWFQYKILNKSC 112
DB 73 LDSCPPHNPFLKREKLCFSQVYVWPMTNTINLVKESC 110
RESULT 9
UDROP1
cystatin S precursor - human
N;Alternate names: cystatin SA-III; salivary acidic protein-1
C;Species: Homo sapiens (man)
C;Date: 25-Feb-1985 #sequence_revision 08-Feb-1996 #text_change 16-Jul-1999
C;Accession: S17667; S16500; A01272; A29603; S19280; A56508
R;Bobek, L.A.; Aguirre, A.; Levine, M.J.
Biochem. J. 278, 629-635, 1991
A;Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ an
A;Reference number: S17667; MUID:91378918; PMID:1898352
A;Accession: S17667
A;Molecule type: mRNA
A;Residues: 1-141 <SOB>
A;Cross-references: EMBL:X54667; NID:G30365; PIDN:CAA38478.1; PID:G30366
R;Lamkin, M.S.; Jensen, J.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.
Arch. Biochem. Biophys. 288, 664-670, 1991
A;Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle,
A;Reference number: S16500; MUID:91378515; PMID:1898055
A;Accession: S16500
A;Status: preliminary
A;Molecule type: protein
A;Residues: 21-134,'D',136-141 <IHU>
R;Isemura, S.; Saitoh, E.; Sanada, K.
J. Biochem. 96, 489-498, 1984
A;Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal
A;Reference number: A91985; MUID:85054716; PMID:6501254

Db 80 LTNCPEHDPHLMRKALCSFQIYVPMWKGTHTLTKSSCKN 119

RESULT 5

B29632

Cystatin SA precursor - human

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999

C;Accession: B29632; S02490; A41422; B27015

R;Saïto, E.; Kim, H.S.; Smithies, O.; Maeda, N.

Gene 61, 329-338, 1987

A;Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three members

A;Reference number: A91589; MUID:88185836; PMID:3446578

A;Accession: B29632

A;Molecule type: DNA

A;Residues: 1-141 <SAI>

A;Cross-references: GB:M19673; GB:M19170; NID:G186403; PIDN:AAA36116.1; PID:G386826

A;Note: the authors translated the codon GAC for residue 129 as Asn

R;Saïto, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.

Biol. Chem. Hoppe-Seyler 369, 191-197, 1988

A;Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily

A;Reference number: S02489; MUID:89076505; PMID:3202984

A;Accession: S02490

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 21-141 <SA2>

R;Isemura, S.; Saïto, E.; Sanada, K.

J. Biochem. 102, 693-704, 1987

A;Title: Characterization and amino acid sequence of a new acidic cysteine proteinase in

A;Reference number: A41422; MUID:88139220; PMID:3436950

A;Accession: A41422

A;Molecule type: protein

A;Residues: 25-141 <ISE>

R;Isemura, S.; Saïto, E.; Sanada, K.; Ito, S.

In Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruy

A;Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.

A;Reference number: A27015

A;Accession: B27015

A;Molecule type: protein

A;Residues: 25-134, 'D', 136-141 <IS2>

C;Genetics:

A;Gene: GDB:CST2

A;Cross-references: GDB:119816; OMIM:123856

A;Map position: 20p11.2-20p11.2

C;Superfamily: cystatin; cystatin homology

F;30-141/Domain: cystatin homology <CYS>

Query Match 25.5%; Score 160.5; DB 2; Length 141;

Best Local Similarity 31.5%; Pred. No. 6.3e-09;

Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKSDDKYHFRIFVLKQVQVTHLEHVLNVMQWTTCCQK--PETNVCVP 81

Db 47 ALHFVISEYNKATEDEYRLLRLVLRAREQIVGVNFFDIEVGRITCTKSQPNLDTCAF 106

QY 82 QER-ELHKQVNCFSVPFVAFWPEFQYKILNKSC 112

Db 107 HEQPELQKQLCSFQIYEVPWEDRMSLVNSRC 138

RESULT 6

A36163

Cystatin C precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 16-Jul-1999

C;Accession: A36163

R;Solem, M.; Rawson, C.; Lindburg, K.; Barnes, D.

Biochem. Biophys. Res. Commun. 172, 945-951, 1990

A;Title: Transforming growth factor beta regulates cystatin C in serum-free mouse embry

A;Reference number: A36163; MUID:91054522; PMID:2241983

A;Accession: A36163

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-140 <SOL>

A;Cross-references: EMBL:M59470; NID:G192911; PIDN:AAA63298.1; PID:G192912

C;Superfamily: cystatin; cystatin homology

F;29-140/Domain: cystatin homology <CYS>

F;93-103,117-137/Diulfide bonds: #status predicted

Query Match 24.9%; Score 156.5; DB 2; Length 140;

Best Local Similarity 31.4%; Pred. No. 1.6e-08;

Matches 33; Conservative 24; Mismatches 45; Indels 3; Gaps 2;

QY 13 EWMVAVNYAKDSLOWITDQYNKSDDKYHFRIFVLKQVQVTHLEHVLNVMQWTTCCQ 72

Db 35 EADANEQVRRALDFAVSEYKNGSDAYHSRAIQVVRARQLVAGVNYFPDVMGRITCT 94

QY 73 KPET--TNC-VPOERELHKQVNCFFSVFVAFWPEFQYKILNKSCSS 114

Db 95 KSQTNLTDCPFDPHLMRKALCSFQIYVPMWKGTHSLTKFSCKN 139

RESULT 7

UDHU

Cystatin C precursor [validated] - human

N;Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma

C;Species: Homo sapiens (man)

C;Date: 06-Jul-1982 #sequence_revision 31-Mar-1991 #text_change 08-Dec-2000

C;Accession: S10216; S00004; A33400; S02751; A01370; A25434; S12288; A32732; A60

R;Abrahamson, M.; Olafsson, I.; Palsdottir, A.; Ulvsbaeck, M.; Lundwall, A.; Jensen, O.

Biochem. J. 268, 287-294, 1990

A;Title: Structure and expression of the human cystatin C gene.

A;Reference number: S10216; MUID:90303202; PMID:2363674

A;Accession: S10216

A;Molecule type: DNA

A;Residues: 1-146 <AB1>

A;Cross-references: EMBL:X52255; NID:G30257; PIDN:CAA36497.1; PID:G296643

R;Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FEBS Lett. 216, 229-233, 1987

A;Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A;Reference number: S00004; MUID:87219149; PMID:3495457

A;Accession: S00004

A;Molecule type: mRNA

A;Residues: 1-146 <AB2>

A;Cross-references: EMBL:X05607; NID:G30371; PIDN:CAA29096.1; PID:G755738

R;Levy, E.; Lopez-Otin, C.; Ghiso, J.; Geltner, D.; Frangione, B.

J. Exp. Med. 169, 1771-1778, 1989

A;Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A;Reference number: JLO095; MUID:89235594; PMID:2541223

A;Accession: JLO095

A;Molecule type: DNA

A;Residues: 1-146 <LEV>

A;Cross-references: GB:X61681; NID:G30367; PIDN:CAA43856.2; PID:G4490944

A;Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit

R;Saïto, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, I.

Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989

A;Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which

A;Reference number: A33400; MUID:89350949; PMID:2764935

A;Accession: A33400

A;Molecule type: DNA

A;Residues: 1-24, 'T', 26-146 <SAI>

A;Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:G181385; PIDN:AAA52164.1; PID:G

R;Ghiso, J.; Cowan, N.; Frangione, B.

Biol. Chem. Hoppe-Seyler 369, 205-208, 1988

A;Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron

A;Reference number: S02751; MUID:89076507; PMID:3264504

A;Accession: S02751

A;Molecule type: DNA

A;Residues: 82-119 <GH2>

A;Cross-references: EMBL:M27769

A;Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown

R;Grubb, A.; Lofberg, H.

Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982

A;Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in th

A;Reference number: A01270; MUID:82222268; PMID:6283552

A:Residues: 24-139 <RUE>
R;Laber, B.; Krigelsstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.
FEBS Lett. 248, 162-168, 1989
A:Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.
A:Reference number: S04008; MUID:89252033; PMID:2721673
A:Accession: S04008
A:Molecule type: protein
A:Residues: 97-114 <LAB>
R;Colella, R.; Bird, J.W.C.
Gene 130, 175-181, 1993
A:Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping t:
A:Reference number: JN0789; MUID:93366172; PMID:8359684
A:Accession: JN0789
A:Molecule type: DNA
A:Residues: 1-139 <CO2>
A:Cross-references: GB:M95725
A:Note: authors failed to translate the codon for residue 115-Tyr
C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases;
C:Genetics:
A:Gene: Csn
A:Introns: 76/3; 114/3
C:Superfamily: cystatin; cystatin homology
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-139/Product: cystatin, long form #status experimental <CYLP>
F:30-139/Domain: cystatin homology <CYS>
F:32-139/Product: cystatin, short form #status experimental <CYSF>
F:76-80/Region: inhibitory #status predicted
F:94-104,118-138/Diulfide bonds: #status experimental
F:103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 27.1%; Score 170.5; DB 1; Length 139;
Best Local Similarity 37.0%; Pred. No. 6.3e-10;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

Qy 24 SLOWITDQYNKESDDKYHFRVLKVRQVTDHLEYHLNVEMQWTTCKP--ETTNC-V 80
Db 47 ALQFAMAEYNRNSNDKYSRRVVRVISAKQLVSGIKYILQVIGRTTCPSKSGDLSCEF 106
Qy 81 PQERELHKQNCFFSFVAPVPWFQYKILNKSC 112
Db 107 HDEPMAYKTYCTFVVYSIFWLNQIKLESKC 138

RESULT 4
S10587
Cystatin C - rat
C:Species: Rattus sp. (rat)
C:Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C:Accession: S10587
R;Esnard, F.; Esnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gauthi:
Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990
A:Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat
A:Reference number: S10587; MUID:90380276; PMID:2400577
A:Accession: S10587
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-120 <ESN>
A:Note: 43-Asn was also found
A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala
C:Superfamily: cystatin; cystatin homology
F:9-120/Domain: cystatin homology <CYS>

Query Match 26.6%; Score 167.5; DB 2; Length 120;
Best Local Similarity 33.0%; Pred. No. 1.1e-09;
Matches 33; Conservative 24; Mismatches 40; Indels 3; Gaps 2;

Qy 18 ENYAKDSLQWITDQYNKESDDKYHFRVLKVRQVTDHLEYHLNVEMQWTTCKPET- 76
Db 20 EGVQRALDFAVSEYNGKSNDAVHSRAIQVVPARQLVAGINYYLDVEMGRITCTKSQTN 79
Qy 77 -TNC-VPQERELHKQVNCFFSFVAPVPWFQYKILNKSCSS 114

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 22.1329 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-3
Perfect score: 629
Sequence: 1 YQARKKTFLSVHEWAVENY.....VFAVPWFQYKILNKSCSD 115

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl:*
2: Pirl:*
3: Pirl:*
4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	34.0	139	2 A45361	cystatin-related e
2	171.5	27.3	127	2 S07085	cystatin C precurs
3	170.5	27.1	139	1 UDCH	cystatin precursor
4	167.5	26.6	120	2 S10587	cystatin C - rat
5	160.5	25.5	141	2 B29632	cystatin SA precur
6	156.5	24.9	140	2 A36163	cystatin C precurs
7	155	24.6	146	1 UDHU	cystatin C precurs
8	153.5	24.4	112	1 UDHO	cystatin - bovine
9	146.5	23.3	141	1 UDHUP1	cystatin S precurs
10	139.5	22.2	141	1 UDHUP2	cystatin SN precur
11	134	21.3	111	1 JC2040	cystatin - chum sa
12	131.5	20.9	142	2 A47442	cystatin D precurs
13	129.5	20.6	141	2 JQ1470	cystatin S precurs
14	127	20.2	111	2 A28793	cystatin - puff ad
15	126	20.0	132	2 JC4918	cystatin precursor
16	102.5	16.3	133	2 JC4536	cystatin precursor
17	98.5	15.7	434	1 KGBOL2	kininogen, LMW II
18	98.5	15.7	619	1 KGBOH2	kininogen, HMW II
19	96.5	15.3	436	1 KGBOL1	kininogen, LMW I p
20	96.5	15.3	621	1 KGBOH1	kininogen, HMW I p
21	95.5	15.2	162	2 A43428	onchocystatin - ne
22	93.5	14.9	427	1 KGHUL1	kininogen, LMW pre
23	93.5	14.9	644	1 KGHUL1	kininogen, HMW pre
24	87	13.8	433	2 A28055	K-kininogen, LMW I
25	87	13.8	639	2 A25486	kininogen, HMW I p
26	81	12.9	430	2 A23897	major acute phase
27	81	12.9	430	2 B28055	T-kininogen, LMW I
28	80	12.7	498	2 T31871	hypothetical prote
29	79	12.6	430	1 KGRITI	T-kininogen I prec

30	77	12.2	430	2 D64151	hypothetical prote
31	75	11.9	400	2 A46297	beta-1,6-N-acetylgl
32	75	11.9	423	1 KGRIM	major acute phase
33	74	11.8	582	2 S43613	membrane protein p
34	73.5	11.7	438	2 A47702	glucan 1,3-beta-gl
35	73.5	11.7	438	2 T52149	beta-glucanase lim
36	73	11.6	861	2 S12499	CHLI protein - yea
37	71.5	11.4	132	2 D88508	protein H14A12.5 f
38	71.5	11.4	1227	2 T23004	hypothetical prote
39	71.5	11.4	1779	2 T23130	hypothetical prote
40	71	11.3	677	1 SYECMT	methionine-tRNA 11
41	71	11.3	677	2 H85839	methionine tRNA sy
42	71	11.3	677	2 H90993	methionine tRNA sy
43	70.5	11.2	448	2 JN0118	glucan 1,3-beta-gl
44	70	11.1	184	2 D72061	hypothetical prote
45	70	11.1	184	2 B86562	hypothetical prote

ALIGNMENTS

RESULT 1

A45361
cystatin-related epididymal specific protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A45361
R;Corrwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.
Mol. Endocrinol. 6, 1653-1664, 1992
A;Title: The CRIS gene: a unique testis-regulated gene related to the cystatin family is
A;Reference number: A45361, MUID:93078799; PMID:1280328
A;Accession: A45361
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-139 <OR>
A;Cross-references: GB:S49926; NID:G260492; PIDN:AAC35390.1; PID:G260493
A;Note: sequence extracted from NCBI backbone (NCBIP:118813)
C;Superfamily: cystatin; cystatin hemology
F;28-139/Domain: cystatin homology <Cys>

Query Match	34.0%	Score 214;	DB 2;	Length 139;
Best Local Similarity	37.1%;	Pred. NO. 2.9e-14;		
Matches	43;	Conservative 29;	Mismatches 30;	Indels 14; Gaps 4;
QY	10	SVHEWAVENY-----AKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLE	59	
DB	22	SKNEVKA-QNYFGSINISNANVKQCVWFAMKYNKESDKYFLVDKILHAKLQITDRME	80	
QY	60	YHLNVEMQWTCQK-ETTNCVPQER-ELHKQVNCFFSVFAVPWFPEQYKILNKSC	112	
DB	81	YQIDVQISRNSCKKPLNNTENCIPQKKPELEKMKSCSLVGLGALPWNGEFNLLSKEC	136	

RESULT 2

S07085
cystatin C precursor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999
C;Accession: S07085; S01337; S21109
R;Cole, T.; Dickson, P.W.; Esmard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Schre
Eur. J. Biochem. 186, 35-42, 1989
A;Title: The cDNA structure and expression analysis of the genes for the cysteine protei
A;Reference number: S07085; MUID:90092122; PMID:2869174
A;Accession: S07085
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-127 <COL>
A;Cross-references: EMBL:X16957; NID:G56041; PIDN:CAA34831.1; PID:G736290
R;Esmard, A.; Esmard, F.; Faucher, D.; Gauthier, F.
FEBS Lett. 236, 475-478, 1988
A;Title: Two rat homologues of human cystatin C.
A;Reference number: S01337; MUID:88313020; PMID:3044831
A;Accession: S01337

; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-6

Query Match 30.0%; Score 189; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. NO. 3.6e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 39
DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDK 36

RESULT 14
US-09-941-314-8
; Sequence 8, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-8

Query Match 29.7%; Score 187; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. NO. 5.9e-14;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 KDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTD 56
DB 1 KDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTD 35

RESULT 15
US-09-740-638-2
; Sequence 2, Application US/09740638
; Patent No. US20020006856A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Zcys8: A Member of the Cystatin
; TITLE OF INVENTION: Superfamily
; FILE REFERENCE: 99-104
; CURRENT APPLICATION NUMBER: US/09/740,638
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-740-638-2

Query Match 28.2%; Score 177.5; DB 9; Length 145;

Best Local Similarity 31.4%; Pred. No. 4e-12;
Matches 32; Conservative 30; Mismatches 37; Indels 3; Gaps 2;
QY 13 EYMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCC 72
DB 34 KLMSKKK-MNSTLNFFIQSYNNASNDTYLYRVQRLIRSORQLTGVGEYIVTVKIGWTKCK 92
QY 73 KPETT--CVPORELHKQVNCFFSVFAVPWFQYKILNKSC 112
DB 93 RNDTSNSSCPLOSKKURKSLICESLIYTWPIWNYFQLWNNSC 134

Search completed: March 18, 2004, 14:30:52
Job time : 70.5933 secs

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53
 Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 10
 US-09-864-761-48936
 ; Sequence 48936, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864, 761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 48936
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL096677.18
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
 ; OTHER INFORMATION: EST HUMAN HIT: AI200857.1, EVALUATE 5.00e-23
 ; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
 US-09-864-761-48936

Query Match 41.7%; Score 262; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53
 Db 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 11
 US-09-941-314-10
 ; Sequence 10, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR APPLICATION NUMBER: 60/230,230
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 46
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-10

Query Match 40.4%; Score 254; DB 9; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 78
 Db 1 NKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCKPPTTN 46

RESULT 12
 US-09-941-314-12
 ; Sequence 12, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
 ; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
 ; FILE REFERENCE: 00-81PC
 ; CURRENT APPLICATION NUMBER: US/09/941,314
 ; CURRENT FILING DATE: 2001-08-29
 ; PRIOR APPLICATION NUMBER: 60/230,230
 ; PRIOR FILING DATE: 2001-09-01
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 33
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-941-314-12

Query Match 30.0%; Score 189; DB 9; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.3e-14;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 QVTDHLEYHLNVEMQWTTCKPPTTNQVQERE 85
 Db 1 QVTDHLEYHLNVEMQWTTCKPPTTNQVQERE 33

RESULT 13
 US-09-941-314-6
 ; Sequence 6, Application US/09941314
 ; Patent No. US20020142396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZymoGenetics, Inc.
 ; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to

```
; ORGANISM: Homo sapiens
US-09-941-314-14
Query Match      45.8%; Score 288; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 3.4e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 KESDDKYHFRIFVLKQVQVTDHLEHLNVNQWTTCKPPTNCVPOERE 85
      |||
      1 KESDDKYHFRIFVLKQVQVTDHLEHLNVNQWTTCKPPTNCVPOERE 52

RESULT 7
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-17

Query Match      43.4%; Score 273; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 NVEMQWTTCKPPTNCVPOERHLKQVNCFFSVFVPEQYKILNK 110
      |||
      1 NVEMQWTTCKPPTNCVPOERHLKQVNCFFSVFVPEQYKILNK 48

RESULT 8
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-13

Query Match      43.2%; Score 272; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 DQYNKESDDKYHFRIFVLKQVQVTDHLEHLNVNQWTTCKPPTTN 78
      |||
      1 DQYNKESDDKYHFRIFVLKQVQVTDHLEHLNVNQWTTCKPPTTN 49

RESULT 9
Query Match      41.7%; Score 262; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;

US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34822
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109954.10
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match      41.7%; Score 262; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
```

```
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-4

Query Match          100.0%; Score 629; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.1e-63;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFLSVHVEVAVENTYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 60
DB 3 YQARKKTFLSVHVEVAVENTYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 62

QY 61 HLNVEQWTTTCQKPEITNCVQPERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115
DB 63 HLNVEQWTTTCQKPEITNCVQPERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 117

RESULT 3
US-09-941-314-2
; Sequence 2, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-2

Query Match          100.0%; Score 629; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.6e-63;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFLSVHVEVAVENTYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 60
DB 23 YQARKKTFLSVHVEVAVENTYAKDSLQWITDQYNKESDDKYHFRIFRLVKVQVQVTDHLEY 82

QY 61 HLNVEQWTTTCQKPEITNCVQPERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115
DB 83 HLNVEQWTTTCQKPEITNCVQPERLHKQVNCFFSVFVAVPWFQYKILNKSCSSD 137

RESULT 4
US-09-941-314-15
; Sequence 15, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
```

```
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-15

Query Match          70.9%; Score 446; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.3e-43;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 QYNKESDDKYHFRIFRLVKVQVQVTDHLEYHLNVEMQWTTTCQKPEITNCVQPERLHKQV 90
DB 1 QYNKESDDKYHFRIFRLVKVQVQVTDHLEYHLNVEMQWTTTCQKPEITNCVQPERLHKQV 60

QY 91 NCFFSVFVAVPWFQYKILNK 110
DB 61 NCFFSVFVAVPWFQYKILNK 80

RESULT 5
US-09-941-314-16
; Sequence 16, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-16

Query Match          53.1%; Score 334; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 RQVTDHLEYHLNVEMQWTTTCQKPEITNCVQPERLHKQVNCFFSVFVAVPWFQYKILNK 110
DB 1 RQVTDHLEYHLNVEMQWTTTCQKPEITNCVQPERLHKQVNCFFSVFVAVPWFQYKILNK 59

RESULT 6
US-09-941-314-14
; Sequence 14, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 69.5933 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-3
Perfect score: 629
Sequence: 1 YQARKKTFSLVHEVMAVENY.....VFAPWFQYKILNKSCSSD 115

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap:
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap:
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap:
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap:
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap:
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap:
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap:
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	629	100.0	115	9 US-09-941-314-3	Sequence 3, Appli
2	629	100.0	117	9 US-09-941-314-4	Sequence 4, Appli
3	629	100.0	117	9 US-09-941-314-2	Sequence 2, Appli
4	446	70.9	80	9 US-09-941-314-15	Sequence 15, Appl
5	334	53.1	59	9 US-09-941-314-16	Sequence 16, Appl
6	288	45.8	52	9 US-09-941-314-14	Sequence 14, Appl
7	273	43.2	48	9 US-09-941-314-17	Sequence 17, Appl
8	272	43.2	49	9 US-09-941-314-13	Sequence 13, Appl
9	262	41.7	50	9 US-09-864-761-34822	Sequence 34822, A
10	262	41.7	50	9 US-09-864-761-48936	Sequence 48936, A
11	254	40.4	46	9 US-09-941-314-10	Sequence 10, Appl
12	189	30.0	33	9 US-09-941-314-12	Sequence 12, Appl
13	189	30.0	36	9 US-09-941-314-6	Sequence 6, Appli
14	187	29.7	35	9 US-09-941-314-8	Sequence 8, Appli
15	177.5	28.2	145	9 US-09-740-638-2	Sequence 2, Appli

16	177.5	28.2	145	13	US-10-006-467-2	Sequence 2, Appli
17	177.5	28.2	145	14	US-10-235-148-2	Sequence 2, Appli
18	171.5	27.3	127	8	US-08-849-303-19	Sequence 19, Appl
19	171.5	27.3	145	14	US-10-168-425-14	Sequence 14, Appl
20	170.5	27.1	116	9	US-09-775-932-16	Sequence 16, Appl
21	170.5	27.1	139	8	US-08-849-303-15	Sequence 15, Appl
22	170.5	27.1	139	9	US-09-969-834-4	Sequence 4, Appli
23	160.5	25.5	121	9	US-09-775-932-8	Sequence 8, Appli
24	160.5	25.5	140	14	US-10-376-564-46	Sequence 46, Appl
25	160.5	25.5	141	8	US-08-849-303-24	Sequence 24, Appl
26	160.5	25.5	141	9	US-09-940-497-6	Sequence 6, Appli
27	159.5	25.4	181	15	US-10-264-049-2608	Sequence 2608, Ap
28	157.5	25.0	165	9	US-09-740-638-5	Sequence 5, Appli
29	157.5	25.0	165	13	US-10-006-467-5	Sequence 5, Appli
30	157.5	25.0	165	14	US-10-235-148-5	Sequence 5, Appli
31	156.5	24.9	140	8	US-08-849-303-18	Sequence 18, Appl
32	156.5	24.9	140	14	US-10-376-564-48	Sequence 48, Appl
33	155	24.6	120	9	US-09-775-932-2	Sequence 2, Appli
34	155	24.6	146	8	US-08-849-303-17	Sequence 17, Appl
35	155	24.6	146	9	US-09-940-497-3	Sequence 3, Appli
36	155	24.6	146	9	US-09-969-834-3	Sequence 3, Appli
37	155	24.6	146	14	US-10-329-428-3	Sequence 3, Appli
38	155	24.6	146	14	US-10-376-564-47	Sequence 47, Appl
39	153.5	24.4	112	8	US-08-849-303-16	Sequence 16, Appl
40	153.5	24.4	118	9	US-09-775-932-24	Sequence 24, Appl
41	146.5	23.3	121	9	US-09-775-932-4	Sequence 4, Appli
42	146.5	23.3	141	8	US-08-849-303-22	Sequence 22, Appl
43	146.5	23.3	141	9	US-09-940-497-5	Sequence 5, Appli
44	146.5	23.3	141	9	US-09-974-298-141	Sequence 141, App
45	146.5	23.3	141	14	US-10-241-220-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1
US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE OF INVENTION: Inhibit Cancer Procoagulant Protein
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-3

Query Match 100.0%; Score 629; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 2.1e-63;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKSDDKYHFRIFRVLKVQROVTDHLEY 60
Db 1 YQARKKTFSLVHEVMAVENYAKDSLOWITDQYNKSDDKYHFRIFRVLKVQROVTDHLEY 60
Qy 61 HLNVENQWTTCCQPEITNCVQPERLHKQVNCFFSVFAVFWFQYKILNKSCSSD 115
Db 61 HLNVENQWTTCCQPEITNCVQPERLHKQVNCFFSVFAVFWFQYKILNKSCSSD 115

RESULT 2
US-09-941-314-4
; Sequence 4, Application US/09941314
; Patent No. US20020142396A1

```

; TELEFAX: 301 309 8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Cystatin SN
; US-08-744-138-6

Query Match      25.5%; Score 160.5; DB 3; Length 141;
Best Local Similarity 31.5%; Pred. No. 6.2e-12;
Matches 29; Conservative 24; Mismatches 36; Indels 3; Gaps 2;

QY      24 SLOWITDQYNKESDDKYHFRIFRVLKVQRQVTDHLEYHLNVEMQWTTCK--PETTNCVP 81
Db      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
47 ALHFVISEYNKATEDEYRRLLEVLRAREQIVGGVNYFFDIEVGRITCKSQPNLDTCAF 106

QY      82 QER-ELHKQVNCFFSVFAVPWFPEQYKILNKSC 112
Db      |:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
107 HEQPELQKKQLCSFQIYEVPEWEDRMSLVNSRC 138

```

Search completed: March 18, 2004, 14:25:34
Job time : 30.4345 secs

RESULT 14
US-09-886-319A-46


```
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PP-0193 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 118195
/ US-08-791-522-4

Query Match 27.1%; Score 170.5; DB 2; Length 139;
Best Local Similarity 37.0%; Pred. No. 3.6e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVMQWTTCKP--ETNC-V 80
Db 47 ALQFAMAEYNRASNDKYSRRVVRVISAKRLQVSGIKYLQVEIGRTTCPKSSGDLQSCF 106

QY 81 POERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 107 HDEPEMAKYTTCTFVVYSIPWLNQIKLLEKSC 138

RESULT 10
US-09-314-777-4
/ Sequence 4, Application US/09314777
/ Patent No. 6110686
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Goli, Surya K.
/ TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
/ TITLE OF INVENTION: PROTEIN
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/314,777
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/791,522
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PP-0193 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
```

```
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 118195
/ US-09-314-777-4

Query Match 27.1%; Score 170.5; DB 3; Length 139;
Best Local Similarity 37.0%; Pred. No. 3.6e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVMQWTTCKP--ETNC-V 80
Db 47 ALQFAMAEYNRASNDKYSRRVVRVISAKRLQVSGIKYLQVEIGRTTCPKSSGDLQSCF 106

QY 81 POERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 107 HDEPEMAKYTTCTFVVYSIPWLNQIKLLEKSC 138

RESULT 11
US-08-849-303-15
/ Sequence 15, Application US/08849303
/ Patent No. 6680424
/ GENERAL INFORMATION:
/ APPLICANT: Atkinson, Howard J.
/ APPLICANT: McPherson, Michael J.
/ APPLICANT: Urwin, Peter E.
/ TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
/ NUMBER OF SEQUENCES: 79
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Klauber & Jackson
/ STREET: 411 Hackensack Avenue, 4th Floor
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/849,303
/ FILING DATE: 21-MAY-1997
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jackson Esq., David A.
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 1321-1-003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-487-5800
/ TELEFAX: 201-343-1684
/ TELEX: 133521
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 139 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ US-08-849-303-15

Query Match 27.1%; Score 170.5; DB 4; Length 139;
Best Local Similarity 37.0%; Pred. No. 3.6e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 24 SLOWITDQYNKESDDKYHFRIFRVLKVRQVTDHLEYHLNVMQWTTCKP--ETNC-V 80
Db 47 ALQFAMAEYNRASNDKYSRRVVRVISAKRLQVSGIKYLQVEIGRTTCPKSSGDLQSCF 106

QY 81 POERELHKQVNCFFSVFVFWPEQYKILNKC 112
Db 107 HDEPEMAKYTTCTFVVYSIPWLNQIKLLEKSC 138
```

13 EYMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCC 72
22 EADASEGVQALDPAVSEYNGSNDAYHSRAIQVVRARQQLVAGINYLDVEMGRITCT 81
73 KPET--TNC-VQERELHKQVCFPSVFAVPWFPEQYKILNKSCSS 114
82 KSQTNLTNCPFFHDQPHLMKALCSFQIYSVPWKGTHLTLSCKN 126

RESULT 8
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cystatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match 27.1%; Score 170.5; DB 4; Length 116;
Best Local Similarity 37.0%; Pred. No. 2.9e-13;
Matches 34; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

24 SLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCC 80
24 ALQFAMAENYRASNDKYSSRVVVISAKQLVSGIKYILQVEIGRTTCPSKSDLSQCE 83
81 PQERELHKQVCFPSVFAVPWFPEQYKILNKSC 112
84 HDEPEMAKYTTCTFVVYSIPMLNQLKLSKC 115

RESULT 9
US-08-791-522-4
; Sequence 4, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,522
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

29.8%; Score 187.5; DB 3; Length 141;
Best Local Similarity 39.6%; Pred. No. 3.1e-15;
Matches 38; Conservative 20; Mismatches 35; Indels 3; Gaps 2;

20 YAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQVTDHLEYHLNVEMQWTTCC--PETT 77
43 YVQHALWYAMKEYNKAANDLYNFRVVDILKSQEQITDSLEYLYEVNIARTMCKKIAGDNE 102
78 NCV-PQERELHKQVCFPSVFAVPWFPEQYKILNKSC 112
103 NCLFQDDPKMKVMFCIFIVSSKPKWKFELKMLKKQC 138

RESULT 7
US-08-849-303-19
; Sequence 19, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urwin, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-849-303-19

Query Match 27.3%; Score 171.5; DB 4; Length 127;
Best Local Similarity 33.3%; Pred. No. 2.4e-13;
Matches 35; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-617-302-3

Query Match 34.0%; Score 214; DB 3; Length 142;
Best Local Similarity 37.1%; Pred. No. 1.8e-18;
Matches 43; Conservative 29; Mismatches 30; Indels 14; Gaps 4;
QY 10 SVHEVMAVENY-----AKDSLOWITDOYNKESDDKYHFRIPRLVKVQROVTDHLE 59
DB 25 SKNEVKA-QNYFSINISNANVCQVWFAMKEYNKESDKYFLVDKILHAKLQITDRME 83
QY 60 YHLNVMQWTTCKP--ETTNCVQPER-ELHKQVNCFFSVFVAPWPFQYKILNKSC 112
DB 84 YQIDVQISRSNCKKPLNNTENCIPQKPELEKMKSCSFLVGALPWNGETVMEKCC 139

RESULT 3
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match 30.0%; Score 188.5; DB 3; Length 142;
Best Local Similarity 40.4%; Pred. No. 2.4e-15;
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;
QY 22 KDSLOWITDOYNKESDDKYHFRIPRLVKVQROVTDHLEHLNVMQWTTCKPPTN--C 79
DB 46 KQCLWFAMQYNEKSEDKYFLVVKTLQALQVTLNLEVLIDVEIARSDCRKPLSTNEIC 105
QY 80 VQPER-ELHKQVNCFFSVFVAPWPFQYKILNKSC 112
DB 106 AIQENSKLRKLSCLSGFLVGALPWNGETVMEKCC 139

RESULT 4
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-4

Query Match 30.0%; Score 188.5; DB 3; Length 142;
Best Local Similarity 40.4%; Pred. No. 2.4e-15;
Matches 38; Conservative 23; Mismatches 30; Indels 3; Gaps 2;
QY 22 KDSLOWITDOYNKESDDKYHFRIPRLVKVQROVTDHLEHLNVMQWTTCKPPTN--C 79
DB 46 KQCLWFAMQYNEKSEDKYFLVVKTLQALQVTLNLEVLIDVEIARSDCRKPLSTNEIC 105
QY 80 VQPER-ELHKQVNCFFSVFVAPWPFQYKILNKSC 112
DB 106 AIQENSKLRKLSCLSGFLVGALPWNGETVMEKCC 139

RESULT 5
US-09-431-480-2
; Sequence 2, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-2

Query Match 29.8%; Score 187.5; DB 3; Length 141;
Best Local Similarity 39.6%; Pred. No. 3.1e-15;
Matches 38; Conservative 20; Mismatches 35; Indels 3; Gaps 2;
QY 20 YAKDSLOWITDOYNKESDDKYHFRIPRLVKVQROVTDHLEHLNVMQWTTCKP--PETT 77
DB 43 YVQHALWYAMKEYNKASNDLYNFRVVDILKSQEQITDSLEYLYLEVINIARTWCKIAGDNE 102
QY 78 NCV-PQERELHKQVNCFFSVFVAPWPFQYKILNKSC 112
DB 103 NCLFQDDPKMKMVFICFIVSSKPKMKFLKMLKKQC 138

RESULT 6
US-09-617-302-2
; Sequence 2, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 29,4345 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-3
Perfect score: 629
Sequence: 1 YQARKKTFLSVHEWMAVENY.....VFAPWFEQKILNKSCSSD 115

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	34.0	142	3	US-09-431-480-3
2	214	34.0	142	3	US-09-617-302-3
3	188.5	30.0	142	3	US-09-431-480-4
4	188.5	30.0	142	3	US-09-617-302-4
5	187.5	29.8	141	3	US-09-431-480-2
6	187.5	29.8	141	3	US-09-617-302-2
7	171.5	27.3	127	4	US-08-849-303-19
8	170.5	27.1	116	4	US-09-775-932-16
9	170.5	27.1	139	2	US-08-791-522-4
10	170.5	27.1	139	3	US-09-314-777-4
11	170.5	27.1	139	4	US-08-849-303-15
12	162	25.8	146	6	5432264-6
13	160.5	25.5	121	4	US-09-775-932-8
14	160.5	25.5	140	4	US-09-886-319A-46
15	160.5	25.5	141	3	US-08-744-138-6
16	160.5	25.5	141	4	US-09-241-376-6
17	160.5	25.5	141	4	US-09-940-497-6
18	160.5	25.5	141	4	US-08-849-303-24
19	156.5	24.9	140	3	US-09-431-480-5
20	156.5	24.9	140	3	US-09-617-302-5
21	156.5	24.9	140	4	US-09-886-319A-48
22	156.5	24.9	140	4	US-08-849-303-18
23	155	24.6	120	4	US-09-775-932-2
24	155	24.6	120	6	5432264-4
25	155	24.6	145	2	US-08-832-535-11
26	155	24.6	146	2	US-08-791-522-3
27	155	24.6	146	3	US-08-744-138-3

28	155	24.6	146	3	US-09-019-485-4	Sequence 4, Appli
29	155	24.6	146	3	US-09-314-777-3	Sequence 3, Appli
30	155	24.6	146	3	US-09-431-480-6	Sequence 6, Appli
31	155	24.6	146	3	US-09-617-302-6	Sequence 6, Appli
32	155	24.6	146	4	US-09-241-376-3	Sequence 3, Appli
33	155	24.6	146	4	US-09-528-436B-3	Sequence 3, Appli
34	155	24.6	146	4	US-09-886-319A-47	Sequence 47, Appli
35	155	24.6	146	4	US-09-940-497-3	Sequence 3, Appli
36	155	24.6	146	4	US-09-976-594-37	Sequence 37, Appli
37	155	24.6	146	4	US-08-849-303-17	Sequence 17, Appli
38	155	24.6	146	5	PCT-US95-07135-9	Sequence 9, Appli
39	153.5	24.4	112	4	US-08-849-303-16	Sequence 16, Appli
40	153.5	24.4	118	4	US-09-775-932-24	Sequence 24, Appli
41	146.5	23.3	121	4	US-09-775-932-4	Sequence 4, Appli
42	146.5	23.3	141	3	US-08-744-138-5	Sequence 5, Appli
43	146.5	23.3	141	3	US-09-431-480-11	Sequence 11, Appli
44	146.5	23.3	141	3	US-09-617-302-11	Sequence 11, Appli
45	146.5	23.3	141	4	US-09-241-376-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-431-480-3
; Sequence 3, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-480-3

Query Match	34.0%	Score 214;	DB 3;	Length 142;
Best Local Similarity	37.1%;	Pred. No. 1.8e-18;		
Matches	43;	Conservative	29;	Mismatches 30; Indels 14; Gaps 4;
QY	10	SVHEWMAVENY-----AKDSLOWITDQNKESDDKYHFRIFRLVKVQVTDHLE	59	
DB	25	SKNEVKA-QNYFGSINISNANVKQCVWFAMKEYNKESDKYFLVDKILHAKLQITDRME	83	
QY	60	YHLNVEMQWTCOKP--ETTNCVQPER-ELHKVQNCFFSVFVFWPEQYKILNKSC	112	
DB	84	YQIDVOISRNCKPLNNTENCIPKKPELEKMKSCSFLVGCALPWNQFNLLSKEC	139	

RESULT 2
US-09-617-302-3
; Sequence 3, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts.
 PT
 PT
 XX
 PS Claim 15; SEQ ID NO 21294; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21532-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 50 AA;
 Query Match 41.7%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

Search completed: March 18, 2004, 14:14:56
 Job time : 106.276 secs

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention
 XX
 SQ Sequence 50 AA;
 Query Match 41.7%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 4 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 53
 DB 1 RKKTFLSVHEVMAVENYAKDSLQWITDQYNKESDDKYHFRIFRVLKVQRQ 50

RESULT 15
 AAM67252
 ID AAM67252 standard; protein; 50 AA.
 XX
 AC AAM67252;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27558.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 FN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US0000668.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234587P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 XX
 PS Example 4; SEQ ID NO 27558; 658pp + Sequence Listing; English.
 XX

XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX PI WPI; 2001-496933/54.
 XX DR
 XX XX
 XX PT New spatially-addressable set of single exon nucleic acid probes, useful
 XX PT for measuring gene expression in sample derived from human breast,
 XX PT comprises number of single exon nucleic acid probes.
 XX Claim 27; SEQ ID NO 15357; 327pp + Sequence Listing; English.
 XX PS
 XX CC The invention relates to a spatially-addressable set of single exon
 XX CC nucleic acid probes for measuring gene expression in a sample derived
 XX CC from human breast and BT 474 cells. The method involves contacting the
 XX CC probes with a collection of detectably labelled nucleic acids derived
 XX CC from mRNA of human breast, and then measuring the label bound to each
 XX CC probe of the microarray. The probes are useful for verifying the
 XX CC expression of regions of genomic DNA predicted to encode proteins. They
 XX CC are useful for gene discovery, and for determining predisposition and/or
 XX CC assessing the toxicity of chemical agents on cells. The microarray of
 XX CC this invention presents a far greater diversity of probes for measuring
 XX CC gene expression, with far less bias than expressed sequence tag
 XX CC microarrays. The method is suitable for rapid production of functional
 XX CC information from genomic sequence. The present sequence is a peptide
 XX CC encoded by a single exon nucleic acid probe of the invention. Note: The
 XX CC sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from WIPO
 XX CC at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 50 AA;
 Query Match 41.7%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 RKKTFLSVHVMVAENYAKDSLOWITDQYNKESDDKYHFRFVLKVRQ 53
 Db 1 RKKTFLSVHVMVAENYAKDSLOWITDQYNKESDDKYHFRFVLKVRQ 50
 RESULT 13
 ABB28913
 ID ABB28913 standard; peptide; 50 AA.
 XX AC ABB28913;
 XX DT 01-FEB-2002 (first entry)
 XX DE Peptide #1564 encoded by breast cell single exon nucleic acid probe.
 XX KW Human; microarray; single exon probe; gene expression; breast; disease;
 XX KW cancer.
 XX OS Homo sapiens.
 XX PN WO200157271-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US000662.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-496933/54.
 XX
 XX PT New spatially-addressable set of single exon nucleic acid probes, useful
 XX PT for measuring gene expression in sample derived from human breast,
 XX PT comprises number of single exon nucleic acid probes.
 XX Claim 27; SEQ ID NO 11881; 327pp + Sequence Listing; English.
 XX PS
 XX CC The invention relates to a spatially-addressable set of single exon
 XX CC nucleic acid probes for measuring gene expression in a sample derived
 XX CC from human breast and BT 474 cells. The method involves contacting the
 XX CC probes with a collection of detectably labelled nucleic acids derived
 XX CC from mRNA of human breast, and then measuring the label bound to each
 XX CC probe of the microarray. The probes are useful for verifying the
 XX CC expression of regions of genomic DNA predicted to encode proteins. They
 XX CC are useful for gene discovery, and for determining predisposition and/or
 XX CC assessing the toxicity of chemical agents on cells. The microarray of
 XX CC this invention presents a far greater diversity of probes for measuring
 XX CC gene expression, with far less bias than expressed sequence tag
 XX CC microarrays. The method is suitable for rapid production of functional
 XX CC information from genomic sequence. The present sequence is a peptide
 XX CC encoded by a single exon nucleic acid probe of the invention. Note: The
 XX CC sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from WIPO
 XX CC at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 50 AA;
 Query Match 41.7%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 RKKTFLSVHVMVAENYAKDSLOWITDQYNKESDDKYHFRFVLKVRQ 53
 Db 1 RKKTFLSVHVMVAENYAKDSLOWITDQYNKESDDKYHFRFVLKVRQ 50
 RESULT 14
 ABB19524
 ID ABB19524 standard; protein; 50 AA.
 XX AC ABB19524;
 XX DT 23-JAN-2002 (first entry)
 XX DE Protein #1523 encoded by probe for measuring heart cell gene expression.
 XX KW Human; gene expression; heart; microarray; vascular system;
 XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
 XX KW congenital heart disease.
 XX OS Homo sapiens.
 XX PN WO200157274-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US000666.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX SQ Sequence 48 AA;
 Query Match 43.4%; Score 273; DB 5; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.6e-23;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 NVEMQWTTCKPPTNCVQPERLHKQVNCFFSVFAVPWFQYKILNK 110
 Db 1 NVEMQWTTCKPPTNCVQPERLHKQVNCFFSVFAVPWFQYKILNK 48

RESULT 8
 AAU79863
 ID AAU79863 standard; peptide: 49 AA.
 XX AC AAU79863;
 XX DT 15-JUL-2002 (first entry)
 XX DE Human cystatin-8 (Zcys8) antigenic fragment #11.
 XX KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX OS Homo sapiens.
 XX PN WO200220567-A2.
 XX PD 14-MAR-2002.
 XX PF 29-AUG-2001; 2001WO-US026868.
 XX PR 01-SEP-2000; 2000US-0320230P.
 XX PA (ZYMO) ZYMOGENETICS INC.
 XX PI Holloway JL, Gao Z, Bishop PD;
 XX DR WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 97-98; 100pp; English.
 XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX SQ Sequence 49 AA;
 Query Match 43.2%; Score 272; DB 5; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 DOYNKESDDKYHFRIFRVLKVQVQVTDHLEVHLNVEMQWTTCKPPTN 78
 Db 1 DOYNKESDDKYHFRIFRVLKVQVQVTDHLEVHLNVEMQWTTCKPPTN 49

RESULT 9
 AAM15096
 ID AAM15096 standard; protein: 50 AA.
 XX AC AAM15096;
 XX DT 12-OCT-2001 (first entry)
 XX DE Peptide #1530 encoded by probe for measuring cervical gene expression.
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX OS Homo sapiens.
 XX PN WO200157278-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US000670.
 XX PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.

XX Claim 27; SEQ ID NO 19922; 487pp; English.
 XX The present invention relates to human single exon nucleic acid probes
 CC (SENPs; see AAI10068-AAI28459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 50 AA;
 Query Match 41.7%; Score 262; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-22;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQV 53
 Db 1 RKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKVQV 50

CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 59 AA;

Query Match 53.1%; Score 334; DB 5; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.3e-30;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 RQVTDHLEYHLNEMQWTTCKPETTNCVPOERLHKQVNCFFSFAVPWFQYKILNK 110
Db 1 RQVTDHLEYHLNEMQWTTCKPETTNCVPOERLHKQVNCFFSFAVPWFQYKILNK 59

RESULT 6
AAU79864
ID AAU79864 standard; peptide; 52 AA.
AC AAU79864;
XX
XX
DT 15-JUL-2002 (first entry)
DE Human cystatin-8 (Zcys8) antigenic fragment #12.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic peptide.
XX
XX Homo sapiens.
XX
XX WO200220567-A2.
XX 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US026868.
XX
XX 01-SEP-2000; 2000US-0230230P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Holloway JL, Gao Z, Bishop PD;
XX WPI; 2002-383044/41.
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
XX Claim 2; Page 98; 100pp; English.
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 59 AA;

CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 52 AA;

Query Match 45.8%; Score 288; DB 5; Length 52;
Best Local Similarity 100.0%; Pred. No. 3.4e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 KESDDKYHFRIFRVLKQVQRTDHLLEYHLNEMQWTTCKPETTNCVPOERE 85
Db 1 KESDDKYHFRIFRVLKQVQRTDHLLEYHLNEMQWTTCKPETTNCVPOERE 52

RESULT 7
AAU79867
ID AAU79867 standard; peptide; 48 AA.
AC AAU79867;
XX
XX
DT 15-JUL-2002 (first entry)
DE Human cystatin-8 (Zcys8) antigenic fragment #15.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic peptide.
XX
XX Homo sapiens.
XX
XX WO200220567-A2.
XX 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US026868.
XX
XX 01-SEP-2000; 2000US-0230230P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Holloway JL, Gao Z, Bishop PD;
XX WPI; 2002-383044/41.
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
XX Claim 2; Page 99; 100pp; English.
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's

CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 137 AA;

Query Match 100.0%; Score 629; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.6e-63;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 60
 DB 23 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRLVKVQROVTDHLEY 82
 QY 61 HLNVMQWTTCKPPTTNCVQREHLKQVNCFFSVFAVPWFQYKILNKCSSD 115
 DB 83 HLNVMQWTTCKPPTTNCVQREHLKQVNCFFSVFAVPWFQYKILNKCSSD 137

RESULT 4
 AAU79865
 ID AAU79865 standard; peptide; 80 AA.
 AC AAU79865;
 XX
 DT 15-JUL-2002 (first entry)
 DE Human cystatin-8 (Zcys8) antigenic fragment #13.
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 98; 100pp; English.
 XX

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in

CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 80 AA;

Query Match 70.9%; Score 446; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 5.8e-43;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 QYNKESDDKYHFRIFRLVKVQROVTDHLEYHLNVMQWTTCKPPTTNCVQREHLKQV 90
 DB 1 QYNKESDDKYHFRIFRLVKVQROVTDHLEYHLNVMQWTTCKPPTTNCVQREHLKQV 60
 QY 91 NCFFSVFAVPWFQYKILNK 110
 DB 61 NCFFSVFAVPWFQYKILNK 80

RESULT 5
 AAU79866
 ID AAU79866 standard; peptide; 59 AA.
 AC AAU79866;
 XX
 DT 15-JUL-2002 (first entry)
 DE Human cystatin-8 (Zcys8) antigenic fragment #14.
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 99; 100pp; English.
 XX

CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 CC
 XX
 SQ Sequence 115 AA;
 Query Match 100.0%; Score 629; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1.3e-63;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQROVTDHLEY 60
 DB 1 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQROVTDHLEY 60
 QY 61 HLNVENQWTTCCQKPEPTTNCVQERELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115
 DB 61 HLNVENQWTTCCQKPEPTTNCVQERELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115
 RESULT 2
 AAU79854
 ID AAU79854 standard; protein; 117 AA.
 XX
 AC AAU79854;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 DR Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 94-95; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 CC
 XX
 SQ Sequence 117 AA;
 Query Match 100.0%; Score 629; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.3e-63;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQROVTDHLEY 60
 DB 3 YQARKKTFLSVHEVMAVENYAKDSLOWITDQYNKESDDKYHFRIFRVLKQVQROVTDHLEY 62
 QY 61 HLNVENQWTTCCQKPEPTTNCVQERELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 115
 DB 63 HLNVENQWTTCCQKPEPTTNCVQERELHKQVNCFFSVFVAVPWFQYKILNKSCSSD 117
 RESULT 3
 AAU79852
 ID AAU79852 standard; protein; 137 AA.
 XX
 AC AAU79852;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8).
 XX
 KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 DR N-PSDB; ABK49522.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 93-94; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 104.276 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-3
Perfect score: 629
Sequence: 1 YQARKKTFLLSVHEVAVENY.....VFAVPWFQYKILNKSCSSD 115

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	100.0	115	5 AAU79853	Aau79853 Human cys
2	629	100.0	117	5 AAU79854	Aau79854 Human cys
3	629	100.0	137	5 AAU79852	Aau79852 Human cys
4	446	70.9	80	5 AAU79865	Aau79865 Human cys
5	334	53.1	59	5 AAU79866	Aau79866 Human cys
6	288	45.8	52	5 AAU79864	Aau79864 Human cys
7	273	43.4	48	5 AAU79867	Aau79867 Human cys
8	272	43.2	49	5 AAU79863	Aau79863 Human cys
9	262	41.7	50	4 AAM15096	Aam15096 Peptide #
10	262	41.7	50	4 ABB34086	Abb34086 Peptide #
11	262	41.7	50	4 AAM27545	Aam27545 Peptide #
12	262	41.7	50	4 ABB32389	Abb32389 Peptide #
13	262	41.7	50	4 ABB28913	Abb28913 Peptide #
14	262	41.7	50	4 ABB19524	Abb19524 Protein #
15	262	41.7	50	4 AAM67252	Aam67252 Human bon
16	262	41.7	50	4 AAM54871	Aam54871 Human bra
17	262	41.7	50	4 ABG48915	Abg48915 Human liv
18	262	41.7	50	4 AAM02833	Aam02833 Peptide #
19	262	41.7	50	5 ABG36903	Abg36903 Human pep
20	254	40.4	46	5 AAU79860	Aau79860 Human cys
21	214	34.0	142	4 AAE02404	Aae02404 Murine cy
22	214	34.0	142	4 AAE04433	Aae04433 Mouse spe
23	214	34.0	143	6 ADA14374	Ada14374 Mouse spe
24	199	31.6	142	7 ADD46708	Add46708 Rat Prote
25	199	31.6	142	7 ADD46704	Add46704 Rat Prote

ALIGNMENTS

RESULT 1
AAU79853
ID AAU79853-standard; protein; 115 AA.

XX AC AAU79853;
XX
XX AC
XX
XX 15-JUL-2002 (first entry)
XX
XX Human cystatin-8 (Zcys8) antigenic fragment #1.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic fragment.
XX
XX Homo sapiens.
XX
XX WO200220567-A2.
XX
XX 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US026868.
XX
XX 01-SEP-2000; 2000US-0230230P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Holloway JL, Gao Z, Bishop PD;
XX
XX WPI; 2002-383044/41.
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.
XX
XX Claim 2; Page 94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large

Aau79862 Human cys
Aau79856 Human cys
Aaw78260 Fragment
Aaw78258 Fragment
Aae02405 Human cys
Aae04434 Human cys
Ada57231 Human sec
Ada41112 Human sec
Adc74335 Human sec
Add37980 Human sec
Add46706 Human Pro
Add46710 Human Pro
Aay96576 Murine cy
Aae02403
Aae04432 Mouse tes
Aau79858 Human cys
Aaw78259 Fragment
Aae04315 Alternati
Aau76555 Human 2cy
Abg75917 Human cys

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 3.21429 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-9

Perfect score: 143
Sequence: 1 DQYNKSDDKYHFRFRVLEKVRQVTD 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	137	CS11_HUMAN	Q9H112 homo sapien
2	96	67.1	139	CS11_MOUSE	Q9D269 mus musculu
3	77	53.8	142	CS18_HUMAN	O60576 homo sapien
4	76	53.1	142	CS18_MOUSE	P32766 mus musculu
5	70	49.0	148	CYT8_RABIT	P01038 gallus galli
6	66	46.2	139	CYT8_CHICK	P09228 homo sapien
7	66	46.2	141	CYT8_HUMAN	P01034 homo sapien
8	65	45.5	146	CYT8_MACMU	O19093 macaca mula
9	65	45.5	146	CYT8_MACMU	O19093 macaca mula
10	65	45.5	146	CYT8_MACMU	O19093 macaca mula
11	64	44.8	142	CYT8_HUMAN	P01061 colurnix co
12	63	44.1	142	CYT8_RAT	O88966 rattus norv
13	62	43.4	127	CYT8_RAT	P14841 rattus norv
14	62	43.4	140	CYT8_MOUSE	P14860 mus musculu
15	62	43.4	148	CYT8_BOVIN	P01035 bos taurus
16	61	42.7	129	CYT8_BOVIN	P35481 cyprinus ca
17	56	39.2	162	CYT8_MACMU	P22085 cyprinus ca
18	54	37.8	141	CYT8_HUMAN	P01037 homo sapien
19	54	37.8	141	CYT8_HUMAN	P01037 homo sapien
20	53	37.1	141	CYT8_RAT	P19313 rattus norv
21	51	35.7	111	CYT8_HUMAN	P08935 blis ariet
22	51	35.7	165	CYT8_HUMAN	Q9H114 homo sapien
23	49	34.3	337	CATV_NPYMC	O64196 homo sapien
24	49	34.3	525	HRG_HUMAN	P04196 homo sapien
25	49	34.3	1146	YHC3_YEAST	O38742 saccharomyc
26	48	33.6	337	CATV_NPYMC	O38742 saccharomyc
27	48	33.6	806	SUS1_ARATH	P10035 arabidopsis
28	48	33.6	917	NIR2_ARATH	P10035 arabidopsis
29	47.5	33.6	505	VP5_AHSVA	Q02166 altilan hor
30	47	32.9	149	CYT8_HUMAN	Q15828 homo sapien
31	47	32.9	224	HLJ1_YEAST	P48353 saccharomyc
32	47	32.9	602	PGH1_MOUSE	P22437 mus musculu
33	47	32.9	602	PGH1_RAT	Q63921 rattus norv

34	46	32.2	734	1	KK08_YEAST	P36004 saccharomyc
35	45.5	31.8	430	1	ENO_BACSU	P37869 bacillus su
36	45	31.5	73	1	YPO1_NPYLD	P30326 lymantria d
37	45	31.5	130	1	CYT8_ONCKE	O98967 oncothynchu
38	45	31.5	130	1	CYT8_ONCKE	O91195 oncothynchu
39	45	31.5	137	1	CST9_MOUSE	O92016 mus musculu
40	45	31.5	258	1	UL79_HSV6U	P52469 human herpe
41	45	31.5	803	1	SUS1_ALINGL	P49034 alnus gluti
42	45	31.5	805	1	SUS1_TULGE	O41608 tulipa gesn
43	45	31.5	918	1	NIA_CUCMA	P17569 cucurbita m
44	44.5	31.1	859	1	PMS2_MOUSE	P54279 mus musculu
45	44	30.8	132	1	MERR_BACCE	P22853 bacillus ce

ALIGNMENTS

RESULT 1
ID CS11_HUMAN STANDARD; PRT; 137 AA.
AC Q9H112; Q9H113;
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Cystatin 11 precursor.
GN CST11 OR CST8L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.B., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dharm P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Frazer A.A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leivaesltho M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McComachie L.J., McElay K., McMurtry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Pratchallam S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
-1- SUBCELLULAR LOCATION: Secreted (Potential).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q9H112-1; Sequence=Displayed;
Name=2;
IsoId=Q9H112-2; Sequence=VSP_001260;
Note=NO experimental confirmation available;
-1- SIMILARITY: Belongs to the cystatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AL096677; CAC13170.1; -
DR EMBL; AL096677; CAC17423.1; -
DR HSSP; P01038; 1A90.
DR Genew; HGNC:15959; CST11.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy; 1.
DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW Thiol protease inhibitor; Signal; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 137
FT SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 93 101 BY SIMILARITY.
FT DISULFID 114 134 BY SIMILARITY.
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 76 110 Missing (in isoform 2).
FT /FTID=VSP 001260.
SQ SEQUENCE 137 AA; 16375 MM; CS8568C39A585C3B CRC64;

Query Match 100.0%; Score 143; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKHFRIFRVLRVQROVTD 27
DB 52 DQNKESDDKHFRIFRVLRVQROVTD 78

RESULT 2

CS11_MOUSE STANDARD; PRT; 139 AA.

AC Q9D269;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

DE Cystatin 11 precursor.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AK020300; BAB32061.1; -
DR HSSP; P01034; 1G96.
DR MGD; MGI:1925490; Cst11.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; Cy; 1.
DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 28
FT CHAIN 29 139
FT SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 94 102 BY SIMILARITY.
FT DISULFID 115 135 BY SIMILARITY.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 139 AA; 16217 MM; F228D9815FA32640 CRC64;

Query Match 67.1%; Score 96; DB 1; Length 139;
Best Local Similarity 61.5%; Pred. No. 1.5e-07;
Matches 16; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 DQNKESDDKHFRIFRVLRVQROVTD 26
DB 53 DQNKESDDKHFRIFRVLRVQROVTD 78

RESULT 3

CST8_HUMAN STANDARD; PRT; 142 AA.

AC Q60676;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

DE Cystatin-related epididymal spermatozoal protein precursor (Cystatin
DE 8).

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramway H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sefra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suleton J.B.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF059244; AAC14707.1; -;
 CC EMBL: AL109954; CAB64234.1; -;
 CC HSSP: P01034; 1G96
 CC Genew: HSCNC12480; CST8
 CC GO: GO:0004869; P:cysteine protease inhibitor activity; TAS.
 CC InterPro: IPR000010; Cystatin.
 CC Pfam: PF00031; Cystatin; 1.
 CC SMART: SM00043; Cy; 1.
 CC K101 protease inhibitor; Signal; Polymorphism.
 CC FT SIGNAL 1 21 POTENTIAL.
 CC FT CHAIN 22 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 CC FT DISULFID 95 105 BY SIMILARITY.
 CC FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT VARIANT 142 142 A -> P (in dbSNP:1054633).
 CC FT SIGNAL 1 19 /FTID=VAR 014527.
 CC FT SEQUENCE 142 AA; 16275 MW; 9A512757E0F4E0D CRC64;
 CC SO
 CC Query Match 53.8%; Score 77; DB 1; Length 142;
 CC Best Local Similarity 53.8%; Pred. No. 0.00011;
 CC Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
 CC Oy 2 QYNKESDCKHFRIFRVLYKQROYTD 27
 CC Db 55 EYNKESDCKVFLVVKTLQALQVNT 80
 CC
 CC RESULT 4
 CC CST8_MOUSE STANDARD; PRT; 142 AA.
 CC AC P32766; O89102;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-
 CC related epididymal specific protein) (Cystatin 8).
 CC GN CST8 OR CRES.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH, and CD-1;
 RX MEDLINE=9924789; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.;
 RT "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermatogenic gene";
 RL Biochem. J. 340:85-93(1999).
 RN [2]
 RP SEQUENCE OF 4-142 FROM N.A.
 RC TISSUE=Epididymis;
 RX MEDLINE=93078799; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;
 RT "The CRES gene: a unique testis-regulated gene related to the cystatin
 RT family is highly restricted in its expression to the proximal region
 RT of the mouse epididymis";
 RL Mol. Endocrinol. 6:1653-1664(1992).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- INDUCTION: Testicular factors or hormones other than androgens
 CC present in the testicular fluid may be involved in the regulation
 CC of CRES gene expression.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF090503; AAC61754.1; -;
 CC EMBL: AF090691; AAC36316.1; -;
 CC EMBL: S49926; AAC35390.1; -;
 CC PIR: A45361; A45361.
 CC HSSP: P01034; 1G96.
 CC MGD: MGI:107161; Cat8.
 CC InterPro: IPR000010; Cystatin.
 CC Pfam: PF00031; Cystatin; 1.
 CC SMART: SM00043; Cy; 1.
 CC K101 protease inhibitor; Signal.
 CC FT SIGNAL 1 19 POTENTIAL.
 CC FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 CC FT DISULFID 95 105 BY SIMILARITY.
 CC FT CARBOHYD 119 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CONFLICT 4 15 PLMSLILFPIIP -> GTRPQNGESGK (IN REF. 2).
 CC FT SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;
 CC SO
 CC Query Match 53.1%; Score 76; DB 1; Length 142;
 CC Best Local Similarity 50.0%; Pred. No. 0.00016;
 CC Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
 CC Oy 2 QYNKESDCKHFRIFRVLYKQROYTD 27
 CC Db 55 EYNKESDCKVFLVVKTLHAKLQITD 80
 CC
 CC RESULT 5
 CC CYTC_RABIT STANDARD; PRT; 148 AA.
 CC AC O97862;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor.
 GN CSTR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eultheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Japanese white; TISSUE=Bone;
 RX MEDLINE=98424349; PubMed=9753427;
 RA Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,
 RA Kawashima H.;
 RT "Large scale isolation of osteoclast-specific genes by an improved
 RT method involving the preparation of a subtracted cDNA library.";
 RL Genes Cells 3:459-475(1998).
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB009342; BAA75921.1; -
 DR HSSP; P01034; 1696.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 KM Thiol protease inhibitor; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 148 CYSTATIN C.
 FT ACT SITE 39 39 REACTIVE SITE.
 FT SITE 83 87 SECONDARY AREA OF CONTACT.
 FT DISUFID 101 111 BY SIMILARITY.
 FT DISUFID 125 145 BY SIMILARITY.
 SQ SEQUENCE 148 AA; 16346 MW; 1523C83116955B9A CRC64;
 Query Match 49.0%; Score 70; DB 1; Length 148;
 Best Local Similarity 45.8%; Pred. No. 0.0014;
 Matches 11; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 QY 2 QYNKESPDKXFRFRVLKXQROV 25
 DB 61 EYNGSNDRYHSRALQVRRARQI 84
 RESULT 6
 CYT CHICK
 ID CYT_CHICK STANDARD; PRT; 139 AA.
 AC P01038;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin precursor (Egg-white cystatin).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90008873; PubMed=2793849;
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
 RT and tissue distribution.";
 RL J. Biol. Chem. 264:17164-17169 (1989).
 RN [2]
 RP SEQUENCE OF 24-139.

RX MEDLINE=94178305; PubMed=6712597;
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
 RT "Cystatin. Amino acid sequence and possible secondary structure.";
 RL Biochem. J. 217:813-817(1984).
 RN [3]
 RP SEQUENCE OF 24-139.
 RX MEDLINE=94110059; PubMed=6662498;
 RA Turk V., Brzin J., Lonjer M., Ritonja A., Eropkin M., Borchart U.,
 RA Machleidt W.;
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 RT of cystatin from chicken egg white.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496 (1983).
 RN [4]
 RP CHARACTERIZATION OF PROTEIN.
 RX MEDLINE=83256421; PubMed=6409085;
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
 RA Suter D.C., Barrett A.J.;
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
 RT purification from egg white, characterization, and detection in
 RT chicken serum.";
 RL Biochem. J. 211:129-138(1983).
 RN [5]
 RP DISULFIDE BONDS.
 RA Grubb A., Loeffberg H., Barrett A.J.;
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 RT cystatin.";
 RL FEBS Lett. 170:370-374 (1984).
 RN [6]
 RP PHOSPHORYLATION.
 RX MEDLINE=89252033; PubMed=2721673;
 RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
 RA Bode W.;
 RT "The cysteine proteinase inhibitor chicken cystatin is a
 RT phosphoprotein.";
 RL FEBS Lett. 248:162-168(1989).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=89052676; PubMed=3191914;
 RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshkov A.,
 RA Brzin J., Kos J., Turk V.;
 RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
 RT its possible mode of interaction with cysteine proteinases.";
 RL EMBO J. 7:2593-2599(1988).
 RN [8]
 RP STRUCTURE BY NMR.
 RX MEDLINE=94087719; PubMed=8263912;
 RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,
 RA Auerswald E.A., Jeanicke R., Oschkinat H.;
 RT "The structures of native phosphorylated chicken cystatin and of a
 RT recombinant unphosphorylated variant in solution.";
 RL J. Mol. Biol. 234:1048-1059(1993).
 CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of
 CC thiol proteases including ficin, papain, and cathepsins B, C, H,
 CC and L. Although isolated from egg white, it is also present in
 CC serum.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J05077; AAA48744.1; -
 DR PIR; A34456; UDCH.
 DR PDB; 1CEW; 31-JAN-94.
 DR PDB; 1A67; 27-MAY-98.
 DR PDB; 1A90; 17-JUN-98.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.

DR PROSITE: PS00287; CYSTATIN; 1.
 KM Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 139
 FT ACT SITE 32 32 CYSTATIN.
 FT SITE 76 80 REACTIVE SITE.
 FT DISUFID 94 104 SECONDARY AREA OF CONTACT.
 FT DISUFID 118 138
 FT MOD RES 103 103
 FT STRAND 35 36 PHOSPHORYLATION (PARTIAL).
 FT TURN 39 40
 FT HELIX 42 51
 FT TURN 52 52
 FT HELIX 53 56
 FT TURN 57 58
 FT STRAND 63 77
 FT STRAND 81 95
 FT TURN 96 97
 FT TURN 99 100
 FT HELIX 101 108
 FT STRAND 115 125
 FT TURN 126 129
 FT STRAND 130 139
 SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;
 Query Match 46.2%; Score 66; DB 1; Length 139;
 Best Local Similarity 45.8%; Pred. No. 0.0052;
 Matches 11; Conservativity 7; Mismatches 6; Indels 0; Gaps 0;
 QY 2 QYNKESDDKXHFPRFVYAKVOROV 25
 Db 54 EYRASNDKSKRVYVIAKROL 77
 RESULT 7
 ID CYT_HUMAN STANDARD; PRT; 141 AA.
 AC P09228; O9UCQ7;
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin SA precursor (Cystatin S5).
 GN Cyst2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88185836; PubMed=3446578;
 RA Saitoh E., Kim H.-S., Smithies O., Maeda N.;
 RT "Human cysteine-protease inhibitors: nucleotide sequence analysis
 of three members of the cystatin gene family.";
 RL Gene 61:329-338(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Cobley S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levesaieho M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prabhalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symamre N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE OF 21-40.
 RC TISSUE-Saliva;
 RX MEDLINE=92138674; PubMed=1778989;
 RA Isemura S., Saitoh E., Sanada K., Minakata K.;
 RT "Identification of full-sized forms of salivary (S-type) cystatins
 (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of
 cystatin S) in human whole saliva and determination of phosphorylation
 sites of cystatin S.";
 RL J. Biochem. 110:648-654(1991).
 RN [4]
 RP SEQUENCE OF 25-141.
 RX MEDLINE=88139220; PubMed=3436950;
 RA Isemura S., Saitoh E., Sanada K.;
 RT "Characterization and amino acid sequence of a new acidic cysteine
 proteinase inhibitor (cystatin SA) structurally closely related to
 cystatin S, from human whole saliva.";
 RL J. Biochem. 102:693-704(1987).
 RN [5]
 RP PRELIMINARY SEQUENCE OF 25-141.
 RA Isemura S., Saitoh E., Sanada K., Isemura M., Ito S.;
 RT "Characterization and amino acid sequence of a new acidic cysteine
 proteinase inhibitor (cystatin SA) structurally closely related to
 cystatin S, from human whole saliva.";
 RL J. Biochem. 102:693-704(1987).
 RN [6]
 RP SEQUENCE OF 25-141 FROM N.A.
 RX MEDLINE=89076505; PubMed=3202964;
 RA Saitoh E., Isemura S., Sanada K., Kim H.-S., Smithies O., Maeda N.;
 RT "Cystatin superfamily. Evidence that family II cystatin genes are
 evolutionarily related to family III cystatin genes.";
 RL Biol. Chem. Hoppe-Seyler 369:191-197(1988).
 CC -1- FUNCTION: Thiol protease inhibitor.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M19673; AAA36116.1; -
 CC EMBL; M19671; AAA36116.1; JOINED.
 CC EMBL; M19672; AAA36116.1; JOINED.
 CC EMBL; A1591074; CAC94784.1; -
 CC PIR; B29632; B29632.
 CC HSSP; P01034; 1G96.
 CC GeneW; HGNC:2474; CST2.
 CC MM; 123856; -
 CC GO; GO:0004869; F: cysteine protease inhibitor activity; TAS.
 CC InterPro; IPR000010; Cystatin.
 CC Pfam; PF00031; cystatin; 1.
 CC SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 KM Thiol protease inhibitor; Signal; Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 141 CYSTATIN SA.

FT ACT_SITE 32 32 REACTIVE SITE.
 FT SITE 76 80 SECONDARY AREA OF CONTACT.
 FT DISULFID 94 104 BY SIMILARITY.
 FT DISULFID 118 138 BY SIMILARITY.
 SQ SEQUENCE 141 AA; 16445 MM; EB54915B1B977AA2 CRC64;
 Query Match 46.2%; Score 66; DB 1; Length 141;
 Best Local Similarity 41.7%; Pred. No. 0.0052;
 Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
 QY 2 QYNESDCKYHFRFRFLKQVROV 25
 DB 54 EYNKATEDEYRRLRLVLRABEQI 77
 RESULT 8
 CYTC_HUMAN
 ID CYTC_HUMAN STANDARD; PRT; 146 AA.
 AC P01034;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
 DE (Post-gamma-globulin).
 GN CST3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=87219149; PubMed=3495457;
 RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;
 RT "Molecular cloning and sequence analysis of cDNA coding for the
 RT precursor of the human cysteine proteinase inhibitor cystatin C";
 RL FEBS Lett. 216:229-233 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukocyte;
 RX MEDLINE=90303202; PubMed=2363674;
 RA Abrahamson M., Olafsson I., Palsdottir A., Uvaback M., Lundwall A.;
 RT "Structure and expression of the human cystatin C gene";
 RL Biochem. J. 268:287-294 (1990).
 RN [3]
 RP SEQUENCE FROM N.A. (HCMVA VARIANT).
 RC TISSUE=Brain;
 RX MEDLINE=89235594; PubMed=2541223;
 RA Levy E., Lopez-Otin C., Ghiso J., Galtner D., Frangione B.;
 RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is
 RT related to a mutation in the cystatin C gene, an inhibitor of
 RT cysteine proteases";
 RL J. Exp. Med. 169:1771-1778 (1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89350949; PubMed=2764935;
 RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.;
 RA Isemura S., Sanada K.;
 RT "The human cystatin C gene (CST3) is a member of the cystatin gene
 RT family which is localized on chromosome 20";
 RL Biochem. Biophys. Res. Commun. 162:1324-1331 (1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Dickinson D.P., Hewett-Emmett D., Thiess M.;
 RT "Acquisition of complex patterns of differential expression in
 RT epithelial cell populations during the evolution of type 2 cystatin
 RT genes";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Codley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leveasliho M.H., Leverisa M.A., Lloyd C., Lloyd D.W., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McCormack L.J., McLeay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Prachinham S.R., Plumb R.W., Ramsay H.,
 RA Phillips B.J.C.T., Prachinham S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Senta H.K., Showman R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871 (2001).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Sapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tomihyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [8]
 RP SEQUENCE OF 27-146.
 RX MEDLINE=82222268; PubMed=6283552;
 RA Grubb A., Loeffberg H.;
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and
 RT presence in the adenylophophysis";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027 (1982).
 RN [9]
 RP SEQUENCE OF 27-73.
 RX MEDLINE=84110059; PubMed=6662498;
 RA Turk V., Brzin J., Longner M., Ritonja A., Eropkin M., Borchart U.,
 RA Machleidt W.;
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 RT of cystatin from chicken egg white";
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496 (1983).
 RN [10]
 RP SEQUENCE OF 27-76.
 RX MEDLINE=84128015; PubMed=6365094;
 RA Brzin J., Popovic T., Turk V.;
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases";
 RL Biochem. Biophys. Res. Commun. 118:103-109 (1984).
 RN [11]
 RP DISULFIDE BONDS.

RA Grubb A., Loeffberg H., Barrett A.J.;
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 RT cystatin.";
 RL FEBS Lett. 170:370-374(1984).
 RN (12)
 RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.
 RX MEDLINE=21173909; PubMed=11276250;
 RA Janowski R., Kozak M., Jankowska E., Gronka Z., Grubb A.,
 RA Abrahamson M., Jaskolski M.;
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through
 RT three-dimensional domain swapping.";
 RL Nat. Struct. Biol. 8:316-320(2001).
 RN (13)
 RP VARIANT GLN-94.
 RX MEDLINE=23316504; PubMed=1352269;
 RA Abrahamson M., Jonsdottir S., Olafsson I., Jenson O., Grubb A.;
 RT "Hereditary cystatin C amyloid angiopathy: identification of the
 RT disease-causing mutation and specific diagnosis by polymerase chain
 RT reaction based analysis.";
 RL Hum. Genet. 89:377-380(1992).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SUBUNIT: Homodimer.
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,
 CC vas deferens, brain, thymus, and ovary and the lowest in the
 CC submandibular gland.
 CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral
 CC hemorrhage with amyloidosis (HCHWA) [MIM:105150]; also known as
 CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis
 CC Icelandic type. HCHWA is characterized by a thickening of the
 CC cerebral arteries walls with deposition of material with the
 CC characteristics of amyloid.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X05607; CAA29096.1; -;
 DR EMBL: X52255; CAA36497.1; -;
 DR EMBL: M27891; AAA52164.1; -;
 DR EMBL: M27889; AAA52164.1; JOINED.
 DR EMBL: M27890; AAA52164.1; JOINED.
 DR EMBL: X61681; CAA43856.2; -;
 DR EMBL: X61682; CAA43856.2; JOINED.
 DR EMBL: X61683; CAA43856.2; JOINED.
 DR EMBL: AF319564; AAK11570.1; -;
 DR EMBL: AL121894; CAC05424.1; -;
 DR EMBL: BC013083; AAI13083.1; -;
 DR PIR: S10216; UDHU.
 DR PDB: 1G96; 06-APR-01.
 DR Genew; HGNC:2475; CST3.
 DR MIM: 604312; -;
 DR MIM: 105150; -;
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; CY: 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 DR Thiol protease inhibitor; Amyloid; Signal; Disease mutation;
 KW Polymorphism; 3D-structure.
 FT SIGNAL 1 26
 FT CHAIN 27 146
 FT ACT SITE 37 37 CYSTATIN C.
 FT SITE 81 85 REACTIVE SITE.
 FT DISULFID 99 109 SECONDARY AREA OF CONTACT.
 FT DISULFID 123 143
 Query Match 45.5%; Score 65; DB 1; Length 146;

Best Local Similarity 41.7%; Pred. No. 0.0077;
 Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
 QY 2 QYNKESDDKXFRIFRVLKXQROV 25
 DB 59 EYNKASNDWYHSRALGVRRKQI 82
 RESULT 9
 ID CYTC_MACMU STANDARD; PRT; 146 AA.
 AC 019092;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor.
 GN CST3.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP MEDLINE=97054523; PubMed=8898820;
 RA Wei L.H., Walker L.C., Levy E.;
 RT "Cystatin C, Icelandic-like mutation in an animal model of
 RT cerebrovascular beta-amyloidosis.";
 RL Stroke 27:2080-2085(1996).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U51912; AAB64050.1; -;
 DR HSSP: P01034; 1G96.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; CY: 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 KW Thiol protease inhibitor; Amyloid; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 146
 FT ACT SITE 37 37 CYSTATIN C.
 FT SITE 81 85 REACTIVE SITE.
 FT DISULFID 99 109 SECONDARY AREA OF CONTACT.
 FT DISULFID 123 143 BY SIMILARITY.
 SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29D26 CRC64;
 Query Match 45.5%; Score 65; DB 1; Length 146;
 Best Local Similarity 41.7%; Pred. No. 0.0077;
 Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
 QY 2 QYNKESDDKXFRIFRVLKXQROV 25
 DB 59 EYNKASNDWYHSRALGVRRKQI 82
 RESULT 10
 ID CYTC_SAISC STANDARD; PRT; 146 AA.
 AC 019093;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Cystatin C precursor.
GN Cyst.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
RL stroke 27:2080-2085(1996).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52028; AAB64051.1; -.
CC HSSP; P01034; 1G96.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; Cystatin; 1.
CC SMART; SM00043; CY; 1.
CC PROSITE; PS00287; CYSTATIN; 1.
CC Thiol protease inhibitor; Amyloid; Signal.
CC SIGNAL 1 26 BY SIMILARITY.
CC CHAIN 1 26 CYPSTATIN C.
CC ACT_SITE 37 146 REACTIVE SITE.
CC SITE 81 85 SECONDARY AREA OF CONTACT.
CC DISULFID 99 109 BY SIMILARITY.
CC FT DISULFID 123 143 BY SIMILARITY.
CC FT DISULFID 123 143 BY SIMILARITY.
CC RN SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;
CC
CC Query Match 45.5%; Score 65; DB 1; Length 146;
CC Best Local Similarity 41.7%; Pred. No. 0.0077;
CC Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 2 QYNKESDDKXHFRIPLVKVQRCV 25
CC DB 59 EYNKASNDKXSRVAVIISAKQQL 82
CC
CC RESULT 11
CC CYT_COTUA STANDARD; PRT; 116 AA.
CC AC P81061;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Cystatin (Egg-white cystatin).
CC OS Coturnix coturnix japonica (Japanese quail).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC OC Coturnix.
CC OX NCBI_TaxID=9394;
CC RN [1]
CC RP SEQUENCE.
CC RS TISSUE=Egg white;
CC RX MEDLINE=97420480; PubMed=9276465;
CC RA GEDLHARTZ B., Engh R.A., Mentele R., Eckerskorn C., Torguato R.,
CC Wiltman J., Kolb H.J., Machleidt W., Filtz H., Auerwald B.A.;
CC RT "Qual cystatin. Isolation and characterisation of a new member of
CC the cystatin family and its hypothetical interaction with cathepsin
CC B."
CC RL FEBS Lett. 412:551-558(1997).

```

```

CC -1- FUNCTION: This protein binds tightly to and inhibits papain and
CC cathepsin B.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC HSSP; P01038; 1CEW.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; Cystatin; 1.
CC SMART; SM00043; CY; 1.
CC PROSITE; PS00287; CYSTATIN; 1.
CC Thiol protease inhibitor; Phosphorylation.
CC ACT_SITE 9 9 REACTIVE SITE.
CC SITE 53 57 SECONDARY AREA OF CONTACT.
CC FT DISULFID 71 81
CC FT DISULFID 95 115
CC FT MOD_RES 80 80 PHOSPHORYLATION.
CC RN SEQUENCE 116 AA; 13093 MW; 48248621053A2E70 CRC64;
CC
CC Query Match 44.8%; Score 64; DB 1; Length 116;
CC Best Local Similarity 41.7%; Pred. No. 0.0085;
CC Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 2 QYNKESDDKXHFRIPLVKVQRCV 25
CC DB 31 EYNKASNDKXSRVAVIISAKQQL 54
CC
CC RESULT 12
CC CST8 RAT STANDARD; PRT; 142 AA.
CC AC O88969;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
CC 8).
CC GN CST8 OR CR8.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;
CC RX MEDLINE=9247899; PubMed=1022962;
CC RA Cornwall G.A., Hsia N., Sutton H.G.;
CC RT "Structure, alternative splicing and chromosomal localization of the
CC rat cystatin-related epididymal spermatogenic gene."
CC RL Biochem. J. 340:85-93(1999).
CC
CC -1- FUNCTION: Performs a specialized role during sperm development and
CC maturation.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the cystatin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF090692; AAC36317.1; -.
CC HSSP; P01034; 1G96.
CC InterPro; IPR000010; Cystatin.
CC Pfam; PF00031; Cystatin; 1.
CC SMART; SM00043; CY; 1.
CC Thiol protease inhibitor; Signal.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 142 CYPSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
CC PROTEIN.
CC SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
CC FT DISULFID 95 105 BY SIMILARITY.
CC FT DISULFID 119 139 BY SIMILARITY.
CC FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).

```

```

SO SEQUENCE 142 AA; 16246 MW; F8873FAA6BCAB34 CRC64;
Query Match 44.1%; Score 63; DB 1; Length 142;
Best Local Similarity 46.2%; Pred. No. 0.015;
Matches 12; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Cy 2 QYNESDDKHFRIFRVLKVGROVTD 27
: ||| ||| : : |||
Db 55 EYNGSEDDKFLPLDKTILHATLQITD 80

RESULT 13
CYTC_RAT STANDARD; PRT; 127 AA.
ID_CYTC_RAT
PI4841;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Fragment).
GN CST3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
[1]
SEQUENCE FROM N.A.
STRAIN=Buffalo;
MEDLINE=90092122; PubMed=2689174;
Cole T., Dickson P.W., Esnard F., Averill F., Risbridger G.,
Gauthier F., Schneider G.,
"the cDNA structure and expression analysis of the
cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
in rat brain.";
Eur. J. Biochem. 186:35-42(1989).
[2]
SEQUENCE OF 8-127.
MEDLINE=90380276; PubMed=2400577;
Esnard F., Esnard A., Faucher D., Capony J.-P., Derancourt J.,
Brillard M., Gauthier F.;
"Rat cystatin C: the complete amino acid sequence reveals a site for
N-glycosylation.";
Biol. Chem. Hoppe-Seyler 371:161-166(1990).
[3]
SEQUENCE OF 8-49.
MEDLINE=88313020; PubMed=3044831;
Esnard A., Esnard F., Faucher D., Gauthier F.;
"Two rat homologues of human cystatin C.";
FEBS Lett. 236:475-478(1988).
[4]
SEQUENCE OF 8-20.
TISSUE=Sertoli cells;
MEDLINE=92225121; PubMed=1563513;
Esnard A., Esnard F., Gullou F., Gauthier F.;
"Production of the cysteine proteinase inhibitor cystatin C by rat
Sertoli cells.";
FEBS Lett. 300:131-135(1992).
-1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
thought to serve an important physiological role as a local
regulator of this enzyme activity. Known to inhibits cathepsin B,
H, and L.
-1- SIMILARITY: Belongs to the cystatin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
DR EMBL; X16957; CAA34831.1; -
DR PIR; S07085; S07085.
DR PIR; S10587; S10587.

```

```

DR HSPSP; P01034; 1G96.
DR InsePro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM0043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
DR K101 protease inhibitor; Signal.
FT NON TER 1 1
FT SIGNAL 1 1
FT CHAIN 8 127 CYSTATIN C.
FT ACT_SITE 18 18 REACTIVE SITE.
FT SITE 62 66 SECONDARY AREA OF CONTACT.
FT DISULFID 80 90 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT CONFLICT 25 25 A - E (IN REF. 2).
SQ SEQUENCE 127 AA; 14039 MM; 78F70158B7925853 CRC64;

Query Match 43.4%; Score 62; DB 1; Length 127;
Best Local Similarity 41.7%; Pred. No. 0.019;
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 2 QYNKESDDKTHFRIFRYLVKYNQVQ 25
DB 40 EYNGSNDVASHRAIQVVRARKOL 63

RESULT 14
CYTC_MOUSE STANDARD; PRT; 140 AA.
ID AC P21460;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Cystatin 3).
GN CST3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=91054522; PubMed=2241983;
RA Solem M., Rawson C., Lindburg K., Barnes D.;
RT "Transforming growth factor beta regulates cystatin C in serum-free
RT mouse embryo (SME) cells.";
RL Biochem. Biophys. Res. Commun. 172:945-951(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=95137392; PubMed=7835704;
RA Hub C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
RT "Structural organization, expression and chromosomal mapping of the
RT mouse cystatin-C-encoding gene (Cstc).";
RL Gene 152:221-226(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ILS, and ISS;
RX MEDLINE=21363810; PubMed=11471062;
RA Ehlinger M.A., Thompson J., Conroy O., Xu Y., Yang F., Cammiff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs.";
RL Mamm. Genome 12:657-663(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feltingold B.A., Grouse L.H., Derge J.G.,
RA Stausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalao D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M59470; AAA63298.1; -;
 DR EMBL: U10098; AAB41056.1; -;
 DR EMBL: AF483486; AAL90760.1; -;
 DR EMBL: AF483487; AAL90761.1; -;
 DR EMBL: BC002072; AAH02072.1; -;
 DR PIR: A36163; A36163.
 DR HSSP: P01034; 1G96.
 DR MGD: MGI:102519; Cat3.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 KM Thiol protease inhibitor; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 140
 FT ACT_SITE 31 31
 FT SITE 75 79
 FT DISULFID 93 103
 FT DISULFID 117 137
 FT CONFLICT 16 16
 FT CONFLICT 84 84
 SQ SEQUENCE 140 AA; 15531 MW; 3A563406DD58D0F5 CRC64;
 L -> F (IN REF. 1).
 Query Match 43.4%; Score 62; DB 1; Length 140;
 Best Local Similarity 41.7%; Pred. No. 0.021;
 Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
 QY 2 QYNKESDPKTHFRIFVLKVRQV 25
 DB 53 EYKNGSNDAYHSPRAIVVRARKL 76
 RESULT 15
 ID CYTC_BOVIN STANDARD; PRT; 148 AA.
 AC P01035;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).
 GN CST3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.
 RC TISSUE=Cerebrospinal fluid, and Choroid plexus;
 RX MEDLINE=9809419; PubMed=9434110;
 RA Olsson S.-L., Ek B., Wilm M., Broberg S., Raak L., Bjork I.;
 RT "Molecular cloning and N-terminal analysis of bovine cystatin C
 RT identification of a full-length N-terminal region.";
 RL Biochem. Biophys. Acta 1343:203-210(1997).
 RN [2]
 RP SEQUENCE OF 37-148.
 RX MEDLINE=85231205; PubMed=3891407;
 RA Hirado M., Tsunawasa S., Sakiyama F., Nitobe M., Fujii S.;
 RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine
 RT proteinase inhibitor.";
 RL FEBS Lett. 186:41-45(1985).
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.
 CC -1- MASS SPECTROMETRY: MW=13420; METHOD=MALDI.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y10811; CAA71771.1; -;
 DR HSSP: P01034; 1G96.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 KM Thiol protease inhibitor; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 30
 FT CHAIN 31 148
 FT MOD_RES 31 31
 FT ACT_SITE 40 40
 FT SITE 84 88
 FT DISULFID 102 112
 FT DISULFID 126 146
 SQ SEQUENCE 148 AA; 16265 MW; EE740FB37CBB9F0E CRC64;
 PROBABLE.
 FT SIGNAL 1 30
 FT CHAIN 31 148
 FT MOD_RES 31 31
 FT ACT_SITE 40 40
 FT SITE 84 88
 FT DISULFID 102 112
 FT DISULFID 126 146
 SQ SEQUENCE 148 AA; 16265 MW; EE740FB37CBB9F0E CRC64;
 PROBABLE.
 Query Match 43.4%; Score 62; DB 1; Length 148;
 Best Local Similarity 41.7%; Pred. No. 0.022;
 Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
 QY 2 QYNKESDPKTHFRIFVLKVRQV 25
 DB 62 EFNKRSNDAYOSRVVRARKV 85
 Search completed: March 18, 2004, 14:16:11
 Job time : 4.21429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 16.1786 Seconds
(without alignments)
526.560 Million cell updates/sec

Title: US-09-941-314-9

Perfect score: 143
Sequence: 1 DQYNKESDDKXHFRIFFVLKVRQVTD 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	138	4	Q8WXU6
2	128	89.5	103	4	Q8WXU5
3	92	64.3	139	11	Q8KSA3
4	81	56.6	141	11	Q9DAP1
5	81	56.6	141	11	Q80ZNS
6	66	46.2	112	13	Q98SR4
7	66	46.2	112	13	Q98SR3
8	64	44.8	81	6	Q29212
9	62	43.4	140	11	Q9BPY9
10	57	39.9	148	5	Q9NH55
11	56	39.2	109	5	Q9TY65
12	56	39.2	161	5	Q16159
13	53	37.8	1320	5	Q81LJ3
14	53	37.8	464	13	Q80125
15	51	35.7	428	10	Q9FYQ9
16	50.5	35.3	394	5	Q9T257

17	50	35.0	108	3	Q9P6S2	Q9P6S2 schizosach
18	50	35.0	319	16	Q9CJ05	Q9CJ05 lactococcus
19	50	35.0	393	10	Q9ZMA1	Q9ZMA1 arabidopsis
20	50	35.0	393	10	Q8GZ44	Q8GZ44 arabidopsis
21	50	35.0	425	3	Q12700	Q12700 debaryomyce
22	49	34.3	205	10	Q946X3	Q946X3 prunus pers
23	49	34.3	341	12	Q8JMC4	Q8JMC4 mamestra co
24	49	34.3	403	5	Q9GPR6	Q9GPR6 dictyosteli
25	49	34.3	442	10	Q9SU04	Q9SU04 arabidopsis
26	49	34.3	498	5	Q16454	Q16454 caenorhabdi
27	49	34.3	2646	12	Q9WP27	Q9WP27 gill-assoct
28	48	33.6	91	5	Q9VHP3	Q9VHP3 drosophila
29	48	33.6	109	16	Q7TU55	Q7TU55 prochloroco
30	48	33.6	140	11	Q80Y72	Q80Y72 mus musculu
31	48	33.6	443	5	Q8ICX8	Q8ICX8 plasmodium
32	48	33.6	522	13	Q7ZU78	Q7ZU78 brachydanio
33	48	33.6	664	17	Q8TKV0	Q8TKV0 methanosaic
34	48	33.6	804	10	Q9AVR8	Q9AVR8 pisum sativ
35	48	33.6	917	10	Q94C76	Q94C76 arabidopsis
36	48	33.6	2133	5	Q813V5	Q813V5 plasmodium
37	47.5	33.2	133	11	Q9D264	Q9D264 mus musculu
38	47.5	33.2	680	10	Q7X6R4	Q7X6R4 oryza sativ
39	47.5	33.2	868	10	Q7XPB6	Q7XPB6 oryza sativ
40	47	32.9	125	5	Q25620	Q25620 onchocerca
41	47	32.9	370	10	Q9M1J6	Q9M1J6 arabidopsis
42	47	32.9	462	13	Q7ZY91	Q7ZY91 xenopus lae
43	47	32.9	462	13	Q7SYH2	Q7SYH2 xenopus lae
44	47	32.9	465	13	Q801B5	Q801B5 xenopus lae
45	47	32.9	606	6	Q97554	Q97554 oryctolagus

ALIGNMENTS

RESULT 1	
Q8WXU6	PRELIMINARY; PRT; 138 AA.
ID Q8WXU6	
AC Q8WXU6	
DT 01-MAR-2002 (TREMBlrel. 20, Created)	
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)	
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)	
DE SC13.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N. A.	
RA Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.;	
RT "SC13: A novel epididymal specific member of the cystatin family.";	
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AF35480; AAL71991.1; -	
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.	
DR InterPro; IPR00010; Cystatin.	
DR Pfam; PF00031; Cystatin; 1.	
DR SMART; SM00043; CY; 1.	
DR SEQUENCE 138 AA; 16506 MW; B49440ACA3585C64 CRC64;	
QY	
Query Match	100.0%; Score 143; DB 4; Length 138;
Best Local Similarity	100.0%; Pred. No. 4,4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	
1 DQYNKESDDKXHFRIFFVLKVRQVTD 27	
53 DQYNKESDDKXHFRIFFVLKVRQVTD 79	
RESULT 2	
Q8WXU5	PRELIMINARY; PRT; 103 AA.
ID Q8WXU5	
AC Q8WXU5	
DT 01-MAR-2002 (TREMBlrel. 20, Created)	
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)	

```

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SC13delta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RI Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335461; AAL71992.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;

Query Match 89.5%; Score 128; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKYHFRIFRYLKVQROVTD 24
Db 53 DQYNKESDDKYHFRIFRYLKVQRO 76

RESULT 3
08K5A3 PRELIMINARY; PRT; 139 AA.
AC 08K5A3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cystatin 11.
GN CST11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA Hamil K.G., Hall S.H.;
RI Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF501290; AAM21709.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 139 AA; 16686 MW; E1E36DB7864D08C CRC64;

Query Match 64.3%; Score 92; DB 11; Length 139;
Best Local Similarity 51.9%; Pred. No. 3e-06;
Matches 14; Conservative 11; Mismatches 2; Indels 0; Gaps 0;

QY 1 DQYNKESDDKYHFRIFRYLKVQROVTD 27
Db 53 EYVNXSEDLNFRILRIKIKEXQMTN 79

RESULT 4
09DAP1 PRELIMINARY; PRT; 141 AA.
AC 09DAP1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 1700006C19R1k protein.
GN 1700006C19R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Saito T., Okazaki Y., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikiado I., Peeole G., Quackenbush J.,
RA Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner U., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guenrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseukl S.,
RA Hasehizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005665; BAB24175.1; -.
DR HSSP; P01038; ICEW.
DR MGD; MGI:1916544; 1700006C19R1k.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16811 MW; C20FA0D8B1AC378C CRC64;

Query Match 56.6%; Score 81; DB 11; Length 141;
Best Local Similarity 53.8%; Pred. No. 0.00015;
Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRIFRYLKVQROVTD 27
Db 54 EYVNXSEDLNFRIVLDIKSQEQITD 79

RESULT 5
080ZNS PRELIMINARY; PRT; 141 AA.
AC 080ZNS;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RIKEN cDNA 1700006C19 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RI Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AAH48681.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; Cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A884951F CRC64;

Query Match 56.6%; Score 81; DB 11; Length 141;
Best Local Similarity 53.8%; Pred. No. 0.00015;
Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRIFRYLKVQROVTD 27

```


Db 54 EYKASNDLYNFRVNDILKSGEOTD 79

RESULT 6

Q96SR4 PRELIMINARY; PRT; 112 AA.

AC Q96SR4; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cystatin (Fragment).
 OS Acipenser sinensis (Chinese sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 NCBI_TaxID=61970;
 RN NCBI_TaxID=61970;
 RP TISSUE=Liver;
 RC TISSUE=Liver;
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two
 species of sturgeons."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF34610; AAK16731.1; -
 DR HSSP; P01038; 1A90.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 FT NON_TER 1
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFED8A08C00 CRC64;

Query Match 46.2%; Score 66; DB 13; Length 112;
 Best Local Similarity 50.0%; Pred. No. 0.024;
 Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKXTHPRIFRYLKYQROV 25

Db 26 EFNKASNDMYIHRVSKVKVQKQV 49

RESULT 7

Q96SR3 PRELIMINARY; PRT; 112 AA.

AC Q96SR3; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cystatin (Fragment).
 OS Acipenser schrenckii (Amur sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 NCBI_TaxID=111304;
 RN NCBI_TaxID=111304;
 RP TISSUE=Liver;
 RC TISSUE=Liver;
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two
 species of sturgeons."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF34611; AAK16732.1; -
 DR HSSP; P01038; 1A90.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 FT NON_TER 1
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFED8A08C00 CRC64;

Query Match 46.2%; Score 66; DB 13; Length 112;

Best Local Similarity 50.0%; Pred. No. 0.024;
 Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 QYNKESDDKXTHPRIFRYLKYQROV 25

Db 26 EFNKASNDMYIHRVSKVKVQKQV 49

RESULT 8

Q29212 PRELIMINARY; PRT; 81 AA.

AC Q29212; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cystatin C precursor (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9623;
 RN NCBI_TaxID=9623;
 RP TISSUE=Small intestine;
 RC MEDLINE=96327607; PubMed=8672129;
 RA Winteroe A.K., Fredholm M., Davies W.;
 RT "Evaluation and characterization of a porcine small intestine CDNA
 library."
 RL Mamm. Genome 7:509-517 (1996).
 DR EMBL; F14794; CAA23262.1; -
 DR HSSP; P01034; 1G96.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 FT NON_TER 81
 SQ SEQUENCE 81 AA; 8771 MW; 25E83E431135B8A7 CRC64;

Query Match 44.8%; Score 64; DB 6; Length 81;
 Best Local Similarity 52.2%; Pred. No. 0.035;
 Matches 12; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 QYNKESDDKXTHPRIFRYLKYQRO 24

Db 59 EYKASXVAXHXXRLVLYXKQ 81

RESULT 9

Q9EPX9 PRELIMINARY; PRT; 140 AA.

AC Q9EPX9; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cystatin C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN NCBI_TaxID=10090;
 RP TISSUE=Small intestine;
 RC STRAIN=BALE/c;
 RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakanson K., Grubb A.,
 RA Gage F.H.;
 RT "FGF-2-Responsive neural stem cell proliferation requires CCG, a novel
 autocrine/paracrine cofactor."
 RT Neuron 28:385-397 (2000).
 RL EMBL; AF311741; AAG40283.1; -
 DR HSSP; P01034; 1G96.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.

FT CHAIN 21 140 CYSTATIN C.
 FT VARIANT 16 A -> G.
 FT SEQUENCE 84 L -> F.
 SO SEQUENCE 140 AA; 15517 MW; 3A563406D58D785 CRC64;

Query Match 43.4%; Score 62; DB 11; Length 140;
 Best Local Similarity 41.7%; Pred. No. 0.12;
 Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 2 QYNKESDDKYHPRFRVLKVKQROV 25
 DB 53 EYKNSNDAYHSPRIQVVPARKOL 76

RESULT 10

ID Q9NH95 PRELIMINARY; PRT; 148 AA.
 AC Q9NH95;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Is-cystatin.
 OS Litomosoides sigmodontis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Litomosoides.
 OX NCBI_TaxID=42156;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;
 RT "Characterization and immunological properties of a cysteine protease
 inhibitor of the filarial parasite Litomosoides sigmodontis."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF229173; AAF35896.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 FT CHAIN 25 148 LS-CYSTATIN.
 SO SEQUENCE 148 AA; 16686 MW; 2950AA89CA5339C9 CRC64;

Query Match 39.9%; Score 57; DB 5; Length 148;
 Best Local Similarity 50.0%; Pred. No. 0.77;
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 4 NKESDDKYHPRFRVLKVKQROV 25
 DB 62 NQOSNDAYHLMPIKLVKSSOV 83

RESULT 11

ID Q9TY65 PRELIMINARY; PRT; 109 AA.
 AC Q9TY65;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Antigen malrose binding protein (Fragment).
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94336252; PubMed=8058358;
 RA Tremblaine K.R., Tree T.I., Gillespie A.J., Guderian R., Maizels R.M.,
 Bradley J.E.;
 RT "Heterogeneity of IgG antibody responses to cloned Onchocerca volvulus
 antigens in microfilaridemia positive individuals from Esmeraldas
 Province, Ecuador."
 RL Parasitol. Immunol. 16:201-209(1994).
 DR EMBL; S71364; AAC60509.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 FT NON TER 1 1
 SO SEQUENCE 109 AA; 12701 MW; 3C6B5EF14B0828B CRC64;

Query Match 39.2%; Score 56; DB 5; Length 109;
 Best Local Similarity 45.5%; Pred. No. 0.8;
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 4 NKESDDKYHPRFRVLKVKQROV 25
 DB 24 NQOSNDAYHLMPIKLVKSSOV 45

RESULT 12

ID O16159 PRELIMINARY; PRT; 161 AA.
 AC O16159;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Cystatin-type cysteine proteinase inhibitor.
 DE BM-CPI-2.
 GN Brugia malayi (filarial nematode worm).
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Brugia.
 OX NCBI_TaxID=6279;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gregory W.F., Blaxter M.L., Maizels R.M.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA Gregory W.F., Maizels R.M.;
 RT "Two distinct cystatin-type cysteine protease inhibitors from the
 parasitic nematode Brugia malayi."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177193; AAB69857.1; -
 DR EMBL; AF177193; AAB51086.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 SO SEQUENCE 161 AA; 18406 MW; 8081351EBE226EB5 CRC64;

Query Match 39.2%; Score 56; DB 5; Length 161;
 Best Local Similarity 45.5%; Pred. No. 1.2;
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OY 4 NKESDDKYHPRFRVLKVKQROV 25
 DB 73 NQOSNDAYHLMPIKLVKSSOV 94

RESULT 13

ID Q81LJ3 PRELIMINARY; PRT; 1320 AA.
 AC Q81LJ3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN PF14_0250.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STAIN=3D7;
 RX MEDLINE=22255705; PubMed=12368864;

RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlson J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguoli S.,
 RA Partee M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
 RA Martin D.M.A., Fairhead A.H., Fraumholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrall B.,
 RT "Genome sequence of the human malaria parasite Plasmodium
 falciparum".
 RL Nature 419:498-511 (2002).
 DR EMBL: AEO14820; AAN36863.1; "-"
 DR GO: GO:0003824; F: catalytic activity; IEA.
 DR GO: GO:0004806; F: triacylglycerol lipase activity; IEA.
 DR GO: GO:0006629; F: lipid metabolism; IEA.
 DR InterPro: IPR002921; Lipase_3.
 DR InterPro: IPR008262; Lipase_AS.
 DR Pfam: PF01764; Lipase_3; 1.
 DR PROSITE: PS00120; LIPASE_SBR; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1320 AA; 155667 MW; 0F06B31D32DF7AF CRC64;

Query Match 37.8%; Score 54; DB 5; Length 1320;
 Best Local Similarity 36.0%; Pred. No. 22;
 Matches 9; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 3 YNKSDDKXHFRIFRVLKVRQVTD 27
 Db 554 YDKREKQYKIAINMLIKYHKELTD 578

RESULT 14
 Q080125 PRELIMINARY; PRT; 464 AA.
 AC 080125;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Fetuin long form.
 OS Cyrtinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OC NCBI_TaxID=7962;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Tsai P.-L., Chang G.-D., Huang C.-J.;
 RT "Purification and cloning of carp fetuin."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY225965; AAC074862.1; "-"
 DR GO: GO:0005874; C: microtubule; IEA.
 DR GO: GO:0004869; F: cysteine protease inhibitor activity; IEA.
 DR GO: GO:0005198; F: structural molecule activity; IEA.
 DR GO: GO:0007018; P: microtubule-based movement; IEA.
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00228; TUBULIN_B AUTOREG; 1.
 SQ SEQUENCE 464 AA; 51698 MW; 7A54F71E44050895 CRC64;

Query Match 37.1%; Score 53; DB 13; Length 464;
 Best Local Similarity 39.1%; Pred. No. 11;
 Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKVRQ 24
 Db 166 KFNKESDHSYFKLMSVGRISTQ 188

RESULT 15
 Q9FYQ9

ID Q9FYQ9 PRELIMINARY; PRT; 428 AA.
 AC Q9FYQ9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Similarity to helicas.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP000421; BAB10021.1; "-"
 DR GO: GO:0004386; F: helicase activity; IEA.
 KW Helicase.
 SQ SEQUENCE 428 AA; 49107 MW; 811FA8B2E790E09 CRC64;

Query Match 35.7%; Score 51; DB 10; Length 428;
 Best Local Similarity 39.1%; Pred. No. 20;
 Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKVRQ 24
 Db 89 QYNVHKEKDHILKFLSTLVQOK 111

Search completed: March 18, 2004, 14:21:28
 Job time : 17.1786 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 41.7103 Seconds

(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254
Sequence: 1 NKESDDDKHFRIRFLVKYQR.....EYHLNVEMQWTTCKPEPTTN 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	46	5	AAU79860 Human cys
2	254	100.0	49	5	AAU79863 Human cys
3	254	100.0	80	5	AAU79865 Human cys
4	254	100.0	115	5	AAU79853 Human cys
5	254	100.0	117	5	AAU79854 Human cys
6	254	100.0	137	5	AAU79852 Human cys
7	248	97.6	52	5	AAU79864 Human cys
8	154	60.6	59	5	AAU79866 Human cys
9	149	58.7	33	5	AAU79862 Human cys
10	138	54.3	24	5	AAU79861 Human cys
11	125	49.2	27	5	AAU79859 Human cys
12	125	49.2	35	5	AAU79858 Human cys
13	118	46.5	92	2	AAW78259 Fragment
14	118	46.5	113	6	ADA57563 Human sec
15	118	46.5	113	6	ADA41457 Human sec
16	118	46.5	113	7	ADC74577 Human sec
17	118	46.5	113	7	ADD38088 Human sec
18	118	46.5	114	2	AAW78153 Human sec
19	118	46.5	123	2	AAW78260 Fragment
20	118	46.5	142	2	AAW78258 Fragment
21	118	46.5	142	4	AAE02405 Human cys
22	118	46.5	142	4	AAE04434 Human cys
23	118	46.5	142	6	ADA57231 Human sec
24	118	46.5	142	6	ADA41112 Human sec
25	118	46.5	142	7	ADC74335 Human sec

26	118	46.5	142	7	ADD37980 Human sec
27	118	46.5	142	7	ADD46706 Human pro
28	118	46.5	142	7	ADD46710 Human pro
29	113	44.5	141	3	AAV96576 Murine cy
30	113	44.5	141	3	AAE02403 Murine cy
31	113	44.5	141	4	AAE04432 Mouse tes
32	111	43.7	142	4	AAE02404 Murine cy
33	111	43.7	142	4	AAE04433 Mouse cys
34	111	43.7	143	6	ADA14374 Mouse spe
35	110	43.3	50	4	AAW15096 Peptide #
36	110	43.3	50	4	ABW34086 Peptide #
37	110	43.3	50	4	AAW27545 Peptide #
38	110	43.3	50	4	ABW32389 Peptide #
39	110	43.3	50	4	ABW28913 Peptide #
40	110	43.3	50	4	ABW19524 Protein #
41	110	43.3	50	4	AAW67252 Human don
42	110	43.3	50	4	AAW54871 Human bra
43	110	43.3	50	4	ABG48915 Human liv
44	110	43.3	50	4	AAW02833 Peptide #
45	110	43.3	50	5	ABG36903 Human pep

ALIGNMENTS

RESULT 1
AAU79860 standard; peptide; 46 AA.
ID AAU79860; AC AAU79860; DT 15-UTL-2002 (first entry)
XX XX
DE Human cystatin-8 (Zcys8) antigenic fragment #8.
XX XX
KW Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis; sperm motility; fertilisation; antigenic peptide.
XX XX
OS Homo sapiens.
XX XX
PN WO200220567-A2.
XX XX
PD 14-MAR-2002.
XX XX
PF 29-AUG-2001; 2001WO-US026868.
XX XX
PR 01-SEP-2000; 2000US-0230230P.
XX XX
PA (ZYMO) ZYMOGENETICS INC.
XX XX
PI Holloway JL, Gao Z, Bishop PD; WPI; 2002-383044/41.
XX XX
PS Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.
XX XX
PS Claim 2; Page 97; 100pp; English.
XX XX
CC The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (1). (1) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(1) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (1) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 46 AA;

Query Match 100.0%; Score 254; DB 5; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.9e-28;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKQROVTDLHYHLNVEQMWTTCQKPEPTN 46
 DB 1 NKESDDKXHFRIFRVLKQROVTDLHYHLNVEQMWTTCQKPEPTN 46

RESULT 2

AAU79863 standard; peptide, 49 AA.

AAU79863;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #11.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 spermatoogenesis, and inhibiting cancer procoagulant protein which leads
 to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97-98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
 polypeptide (1). (1) is useful for: inhibiting cancer procoagulant
 protein in an individual and thus inhibiting the thrombotic events
 associated with cancer; promoting spermatogenesis, modulating seminal
 fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 motility and fertilisation; and as antigenic peptides to generate
 antibodies. Zcys8 is useful as research reagent for characterising sites
 of interaction between Zcys8 and its receptor. Zcys8 is useful in
 enhancing fertilisation during assisted reproduction in humans and in
 animals. Anti-(1) antibodies are useful to screen biological samples like
 blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 presence of Zcys8. The antibodies are also useful to isolate large
 quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 The polynucleotide encoding (1) is useful to detect and to localise the
 expression of a Zcys8 gene in a biological sample and Zcys8
 oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 49 AA;

Query Match 100.0%; Score 254; DB 5; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.1e-28;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKXHFRIFRVLKQROVTDLHYHLNVEQMWTTCQKPEPTN 46
 DB 4 NKESDDKXHFRIFRVLKQROVTDLHYHLNVEQMWTTCQKPEPTN 49

RESULT 3

AAU79865 standard; peptide, 80 AA.

AAU79865;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #13.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 spermatoogenesis; seminal fluid viscosity; cryopreserved sperm;
 sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 spermatoogenesis, and inhibiting cancer procoagulant protein which leads
 to inhibition of thrombotic events associated with cancer.

Claim 2; Page 98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
 polypeptide (1). (1) is useful for: inhibiting cancer procoagulant
 protein in an individual and thus inhibiting the thrombotic events
 associated with cancer; promoting spermatogenesis, modulating seminal
 fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 motility and fertilisation; and as antigenic peptides to generate
 antibodies. Zcys8 is useful as research reagent for characterising sites
 of interaction between Zcys8 and its receptor. Zcys8 is useful in
 enhancing fertilisation during assisted reproduction in humans and in
 animals. Anti-(1) antibodies are useful to screen biological samples like
 blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 presence of Zcys8. The antibodies are also useful to isolate large
 quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 The polynucleotide encoding (1) is useful to detect and to localise the
 expression of a Zcys8 gene in a biological sample and Zcys8
 oligonucleotide probes are useful for in vivo diagnosis. The
 polynucleotide encoding (1) is useful in determining whether a subject's
 chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 copy number changes, insertions, deletions, restriction site changes and
 rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 80 AA;
 Query Match 100.0%; Score 254; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 3.7e-28;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NKESDCKYHFRIRVLKVGROVTDHLEHNLVEMQWTTCKPPTTN 46
 Db 3 NKESDCKYHFRIRVLKVGROVTDHLEHNLVEMQWTTCKPPTTN 48
 RESULT 4
 ID AAU79853 standard; protein; 115 AA.
 AC AAU79853;
 XX
 XX 15-JUL-2002 (first entry)
 DT
 XX Human cystatin-8 (Zcys8) antigenic fragment #1.
 DE
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 PI Holloway JL, Gao Z, Bishop PD;
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 PS Claim 2; Page 94; 100pp; English.
 XX
 XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 115 AA;

Query Match 100.0%; Score 254; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 5.8e-28;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NKESDCKYHFRIRVLKVGROVTDHLEHNLVEMQWTTCKPPTTN 46
 Db 33 NKESDCKYHFRIRVLKVGROVTDHLEHNLVEMQWTTCKPPTTN 78
 RESULT 5
 ID AAU79854 standard; protein; 117 AA.
 AC AAU79854;
 XX
 XX 15-JUL-2002 (first entry)
 DT
 XX Human cystatin-8 (Zcys8) antigenic fragment #2.
 DE
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 PI Holloway JL, Gao Z, Bishop PD;
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 PS Claim 2; Page 94-95; 100pp; English.
 XX
 XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 117 AA;
 Query Match 100.0%; Score 254; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 5.9e-28;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRVLKVGQROVTDHLEHNLVNEQMWTTCQKPEPTN 46
 |||
 DB 35 NKESDDKYHFRIFRVLKVGQROVTDHLEHNLVNEQMWTTCQKPEPTN 80

RESULT 6
 AAU79852
 ID AAU79852 standard; protein; 137 AA.
 AC AAU79852;
 XX
 XX 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8).
 XX
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation.
 XX
 OS Homo sapiens.
 XX
 XX WO200220567-A2.
 PN 14-MAR-2002.
 XX
 XX 29-AUG-2001; 2001WO-US026868.
 PF
 XX 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 DR N-PSDB; ABR49522.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 93-94; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)
 CC
 XX Sequence 137 AA;
 SQ

Query Match 100.0%; Score 254; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 7, 1e-28;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRVLKVGQROVTDHLEHNLVNEQMWTTCQKPEPTN 46

DB 55 NKESDDKYHFRIFRVLKVGQROVTDHLEHNLVNEQMWTTCQKPEPTN 100
 |||
 ID AAU79864 standard; peptide; 52 AA.
 AC AAU79864;
 XX
 XX 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #12.
 XX
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KW sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 XX WO200220567-A2.
 PN 14-MAR-2002.
 XX
 XX 29-AUG-2001; 2001WO-US026868.
 PF
 XX 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 98; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX Sequence 52 AA;
 SQ

Query Match 97.6%; Score 248; DB 5; Length 52;
 Best Local Similarity 100.0%; Pred. No. 1, 6e-27;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KESDDKYHFRIFRVLKVGQROVTDHLEHNLVNEQMWTTCQKPEPTN 46
 |||
 DB 1 KESDDKYHFRIFRVLKVGQROVTDHLEHNLVNEQMWTTCQKPEPTN 45

RESULT 8
AAU79865
ID AAU79866 standard; peptide, 59 AA.
XX
XX AC AAU79866;
XX
XX DT 15-JUL-2002 (first entry)
DE Human cystatin-8 (Zcys8) antigenic fragment #14.
XX
XX CYSTATIN-8; ZCYS8; cancer; procoagulant protein; thrombosis;
KW spermatoogenesis; seminal fluid viscosity; cryopreserved sperm;
KM sperm motility; fertilisation; antigenic peptide.
XX
XX OS Homo sapiens.
PN WO200220567-A2.
PD 14-MAR-2002.
PP 29-AUG-2001; 2001WO-US026868.
PR 01-SEP-2000; 2000US-0230230P.
PS (ZYMO) ZYMOGENETICS INC.
XX
XX PI Holloway JL, Gao Z, Bishop PD;
XX DR WPI; 2002-383044/41.
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
XX to inhibition of thrombotic events associated with cancer.
PS Claim 2; Page 99; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatoogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcys8 is useful as research reagent for characterising sites of interaction between Zcys8 and its receptor. Zcys8 is useful in enhancing fertilisation during assisted reproduction in humans and in CC animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcys8. The antibodies are also useful to isolate large quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes. The polynucleotide encoding (I) is useful to detect and to localise the expression of a Zcys8 gene in a biological sample and Zcys8 oligonucleotide probes are useful for in vivo diagnosis. The polynucleotide encoding (I) is useful in determining whether a subject's chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements and genetic alterations that inactivate the Zcys8 gene. This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 59 AA;

Query Match 60.6%; Score 154; DB 5; Length 59;
Best Local Similarity 100.0%; Pred. No. 4,1e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Cy 20 RQVTDHLEHYHNVMQMWTTCCKPEPTN 46
| | | | | | | | | | | | | | | |
Db 1 RQVTDHLEHYHNVMQMWTTCCKPEPTN 27

RESULT 9
AAU79862
ID AAU79862 standard; peptide, 33 AA.
XX

AC	AAU79862;
XX	
DT	15-JUL-2002 (first entry)
XX	
DE	Human cystatin-B (ZcysB) antigenic fragment #10.
XX	
KW	Cystatin-B; ZcysB; cancer; procoagulant protein; thrombosis; spermatogenesis; seminal fluid viscosity; cryopreserved sperm; sperm motility; fertilisation; antigenic peptide.
XX	
OS	Homo sapiens.
XX	
FN	WO200220567-A2.
PD	14-MAR-2002.
XX	
PF	29-AUG-2001; 2001MO-US026866.
XX	
PR	01-SEP-2000; 2000US-0230230P.
XX	
PA	(ZYMO) ZYMOGENETICS INC.
XX	
PI	Holloway JL, Gao Z, Bishop PD;
XX	
DR	WPI; 2002-383044/41.
PT	Novel isolated mammalian cystatin-B polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.
XX	
PS	Claim 2; Page 97; 100pp; English.
XX	
CC	The invention describes an isolated mammalian cystatin-B (ZcysB) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. ZcysB is useful as research reagent for characterising sites of interaction between ZcysB and its receptor. ZcysB is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of ZcysB. The antibodies are also useful to isolate large quantities of ZcysB protein and DNA sequences that encode ZcysB genes. The polynucleotide encoding (I) is useful to detect and to localise the expression of a ZcysB gene in a biological sample and ZcysB polynucleotide probes are useful for in vivo diagnosis. The polynucleotide encoding (I) is useful in determining whether a subject's chromosomes contain a mutation in the ZcysB gene like aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements and genetic alterations that inactivate the ZcysB gene. This sequence represents an antigenic peptide of human cystatin-B (ZcysB)
XX	
SQ	Sequence 33 AA;
Query Match	58.7%; Score 149; DB 5; Length 33;
Best Local Similarity	100.0%; Pred. NO. 1e-13;
Matches	26; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qv	21 QVTDHLEHYHLNVEVMQMTTCQKPEETT 46 1 QVTDHLEHYHLNVEVMQMTTCQKPEETT 26
Db	
RESULT 10	
ID	AAU79861
XX	AAU79861 standard; peptide; 24 AA.
XX	
DT	15-JUN-2002 (first entry)
XX	

DE Human cystatin-8 (Zcys8) antigenic fragment #9.
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 XX spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.
 XX Homo sapiens.
 OS
 XX MO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 XX 29-AUG-2001; 2001MO-US026668.
 XX
 XX 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 97; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antipodas. Zcys8 is useful as research reagent for characterizing sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 24 AA;
 QY
 Db Query Match 54.3%; Score 138; DB 5; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2.6e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 20 ROYTDHLEYHNVEMQWTCCKPE 43
 1 ROYTDHLEYHNVEMQWTCCKPE 24
 RESULT 11
 AAU79859
 ID AAU79859 standard; peptide; 27 AA.
 AC
 AC AAU79859;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #7.
 XX
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.

KM sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 XX MO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 XX 29-AUG-2001; 2001MO-US026668.
 XX
 XX 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 XX
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 96; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antipodas. Zcys8 is useful as research reagent for characterizing sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 27 AA;
 QY
 Db Query Match 49.2%; Score 125; DB 5; Length 27;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 NKESDDKYHFRIRVYLVKQVQVTD 24
 4 NKESDDKYHFRIRVYLVKQVQVTD 27
 RESULT 12
 AAU79858
 ID AAU79858 standard; peptide; 35 AA.
 AC
 AC AAU79858;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #6.
 XX
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX

PN WO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
XX spermatogenesis, and inhibiting cancer procoagulant protein which leads
XX to inhibition of thrombotic events associated with cancer.
PS
XX Claim 2; Page 96; 100pp; English.
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
XX polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
XX protein in an individual and thus inhibiting the thrombotic events
XX associated with cancer; promoting spermatogenesis, modulating seminal
XX fluid viscosity, enhancing viability of cryopreserved sperm, sperm
XX motility and fertilisation; and as antigenic peptides to generate
XX antibodies. Zcys8 is useful as research reagent for characterising sites
XX of interaction between Zcys8 and its receptor. Zcys8 is useful in
XX enhancing fertilisation during assisted reproduction in humans and in
XX animals. Anti-(I) antibodies are useful to screen biological samples like
XX blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
XX presence of Zcys8. The antibodies are also useful to isolate large
XX quantities of Zcys8 protein and DNA sequences that encode Zcys8 gene.
XX The polynucleotide encoding (I) is useful to detect and to localise the
XX expression of a Zcys8 gene in a biological sample and Zcys8
XX oligonucleotide probes are useful for in vivo diagnosis. The
XX polynucleotide encoding (I) is useful in determining whether a subject's
XX chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
XX copy number changes, insertions, deletions, restriction site changes and
XX rearrangements and genetic alterations that inactivate the Zcys8 gene.
XX This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 35 AA:
Query Match 49.2%; Score 125; DB 5; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKESDDKYHFRIRVLKVRQVTD 24
DB 12 NKESDDKYHFRIRVLKVRQVTD 35
RESULT 13
AAW78259
ID AAW78259 standard; protein; 92 AA.
AC AAW78259;
XX
DT 13-APR-1999 (first entry)
XX
XX Fragment of human secreted protein encoded by gene 28.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischemic shock; Alzheimer's disease; osteoarthritis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.

XX
PN WO9856804-A1.
XX
PD 17-DEC-1998.
XX
XX 11-JUN-1998; 98WO-US012125.
XX
XX 13-JUN-1997; 97US-0049547P.
XX 13-JUN-1997; 97US-0049548P.
XX 13-JUN-1997; 97US-0049549P.
XX 13-JUN-1997; 97US-0049550P.
XX 13-JUN-1997; 97US-0049551P.
XX 13-JUN-1997; 97US-0049552P.
XX 13-JUN-1997; 97US-0049553P.
XX 13-JUN-1997; 97US-0049554P.
XX 13-JUN-1997; 97US-0049555P.
XX 13-JUN-1997; 97US-0049556P.
XX 13-JUN-1997; 97US-0049557P.
XX 13-JUN-1997; 97US-0049558P.
XX 13-JUN-1997; 97US-0049559P.
XX 13-JUN-1997; 97US-0049560P.
XX 13-JUN-1997; 97US-0049561P.
XX 13-JUN-1997; 97US-0049562P.
XX 13-JUN-1997; 97US-0049563P.
XX 13-JUN-1997; 97US-0049564P.
XX 13-JUN-1997; 97US-0049565P.
XX 13-JUN-1997; 97US-0049566P.
XX 13-JUN-1997; 97US-0049567P.
XX 13-JUN-1997; 97US-0049568P.
XX 13-JUN-1997; 97US-0049569P.
XX 13-JUN-1997; 97US-0049570P.
XX 13-JUN-1997; 97US-0049571P.
XX 13-JUN-1997; 97US-0049572P.
XX 13-JUN-1997; 97US-0049573P.
XX 13-JUN-1997; 97US-0049574P.
XX 13-JUN-1997; 97US-0049575P.
XX 13-JUN-1997; 97US-0049576P.
XX 13-JUN-1997; 97US-0049577P.
XX 13-JUN-1997; 97US-0049578P.
XX 13-JUN-1997; 97US-0049579P.
XX 13-JUN-1997; 97US-0049580P.
XX 13-JUN-1997; 97US-0049581P.
XX 13-JUN-1997; 97US-0049582P.
XX 13-JUN-1997; 97US-0049583P.
XX 13-JUN-1997; 97US-0049584P.
XX 13-JUN-1997; 97US-0049585P.
XX 13-JUN-1997; 97US-0049586P.
XX 13-JUN-1997; 97US-0049587P.
XX 13-JUN-1997; 97US-0049588P.
XX 13-JUN-1997; 97US-0049589P.
XX 13-JUN-1997; 97US-0049590P.
XX 13-JUN-1997; 97US-0049591P.
XX 13-JUN-1997; 97US-0049592P.
XX 13-JUN-1997; 97US-0049593P.
XX 13-JUN-1997; 97US-0049594P.
XX 13-JUN-1997; 97US-0049595P.
XX 13-JUN-1997; 97US-0049596P.
XX 13-JUN-1997; 97US-0049597P.
XX 13-JUN-1997; 97US-0049598P.
XX 13-JUN-1997; 97US-0049599P.
XX 13-JUN-1997; 97US-0049600P.
XX 13-JUN-1997; 97US-0049601P.
XX 13-JUN-1997; 97US-0049602P.
XX 13-JUN-1997; 97US-0049603P.
XX 13-JUN-1997; 97US-0049604P.
XX 13-JUN-1997; 97US-0049605P.
XX 13-JUN-1997; 97US-0049606P.
XX 13-JUN-1997; 97US-0049607P.
XX 13-JUN-1997; 97US-0049608P.
XX 13-JUN-1997; 97US-0049609P.
XX 13-JUN-1997; 97US-0049610P.
XX 13-JUN-1997; 97US-0049611P.
XX 13-JUN-1997; 97US-0049612P.
XX 13-JUN-1997; 97US-0049613P.
XX 13-JUN-1997; 97US-0049614P.
XX 13-JUN-1997; 97US-0049615P.
XX 13-JUN-1997; 97US-0049616P.
XX 13-JUN-1997; 97US-0049617P.
XX 13-JUN-1997; 97US-0049618P.
XX 13-JUN-1997; 97US-0049619P.
XX 13-JUN-1997; 97US-0049620P.
XX 13-JUN-1997; 97US-0049621P.
XX 13-JUN-1997; 97US-0049622P.
XX 13-JUN-1997; 97US-0049623P.
XX 13-JUN-1997; 97US-0049624P.
XX 13-JUN-1997; 97US-0049625P.
XX 13-JUN-1997; 97US-0049626P.
XX 13-JUN-1997; 97US-0049627P.
XX 13-JUN-1997; 97US-0049628P.
XX 13-JUN-1997; 97US-0049629P.
XX 13-JUN-1997; 97US-0049630P.
XX 13-JUN-1997; 97US-0049631P.
XX 13-JUN-1997; 97US-0049632P.
XX 13-JUN-1997; 97US-0049633P.
XX 13-JUN-1997; 97US-0049634P.
XX 13-JUN-1997; 97US-0049635P.
XX 13-JUN-1997; 97US-0049636P.
XX 13-JUN-1997; 97US-0049637P.
XX 13-JUN-1997; 97US-0049638P.
XX 13-JUN-1997; 97US-0049639P.
XX 13-JUN-1997; 97US-0049640P.
XX 13-JUN-1997; 97US-0049641P.
XX 13-JUN-1997; 97US-0049642P.
XX 13-JUN-1997; 97US-0049643P.
XX 13-JUN-1997; 97US-0049644P.
XX 13-JUN-1997; 97US-0049645P.
XX 13-JUN-1997; 97US-0049646P.
XX 13-JUN-1997; 97US-0049647P.
XX 13-JUN-1997; 97US-0049648P.
XX 13-JUN-1997; 97US-0049649P.
XX 13-JUN-1997; 97US-0049650P.
XX 13-JUN-1997; 97US-0049651P.
XX 13-JUN-1997; 97US-0049652P.
XX 13-JUN-1997; 97US-0049653P.
XX 13-JUN-1997; 97US-0049654P.
XX 13-JUN-1997; 97US-0049655P.
XX 13-JUN-1997; 97US-0049656P.
XX 13-JUN-1997; 97US-0049657P.
XX 13-JUN-1997; 97US-0049658P.
XX 13-JUN-1997; 97US-0049659P.
XX 13-JUN-1997; 97US-0049660P.
XX 13-JUN-1997; 97US-0049661P.
XX 13-JUN-1997; 97US-0049662P.
XX 13-JUN-1997; 97US-0049663P.
XX 13-JUN-1997; 97US-0049664P.
XX 13-JUN-1997; 97US-0049665P.
XX 13-JUN-1997; 97US-0049666P.
XX 13-JUN-1997; 97US-0049667P.
XX 13-JUN-1997; 97US-0049668P.
XX 13-JUN-1997; 97US-0049669P.
XX 13-JUN-1997; 97US-0049670P.
XX 13-JUN-1997; 97US-0049671P.
XX 13-JUN-1997; 97US-0049672P.
XX 13-JUN-1997; 97US-0049673P.
XX 13-JUN-1997; 97US-0049674P.
XX 13-JUN-1997; 97US-0049675P.
XX 13-JUN-1997; 97US-0049676P.
XX 13-JUN-1997; 97US-0049677P.
XX 13-JUN-1997; 97US-0049678P.
XX 13-JUN-1997; 97US-0049679P.
XX 13-JUN-1997; 97US-0049680P.
XX 13-JUN-1997; 97US-0049681P.
XX 13-JUN-1997; 97US-0049682P.
XX 13-JUN-1997; 97US-0049683P.
XX 13-JUN-1997; 97US-0049684P.
XX 13-JUN-1997; 97US-0049685P.
XX 13-JUN-1997; 97US-0049686P.
XX 13-JUN-1997; 97US-0049687P.
XX 13-JUN-1997; 97US-0049688P.
XX 13-JUN-1997; 97US-0049689P.
XX 13-JUN-1997; 97US-0049690P.
XX 13-JUN-1997; 97US-0049691P.
XX 13-JUN-1997; 97US-0049692P.
XX 13-JUN-1997; 97US-0049693P.
XX 13-JUN-1997; 97US-0049694P.
XX 13-JUN-1997; 97US-0049695P.
XX 13-JUN-1997; 97US-0049696P.
XX 13-JUN-1997; 97US-0049697P.
XX 13-JUN-1997; 97US-0049698P.
XX 13-JUN-1997; 97US-0049699P.
XX 13-JUN-1997; 97US-0049700P.
XX 13-JUN-1997; 97US-0049701P.
XX 13-JUN-1997; 97US-0049702P.
XX 13-JUN-1997; 97US-0049703P.
XX 13-JUN-1997; 97US-0049704P.
XX 13-JUN-1997; 97US-0049705P.
XX 13-JUN-1997; 97US-0049706P.
XX 13-JUN-1997; 97US-0049707P.
XX 13-JUN-1997; 97US-0049708P.
XX 13-JUN-1997; 97US-0049709P.
XX 13-JUN-1997; 97US-0049710P.
XX 13-JUN-1997; 97US-0049711P.
XX 13-JUN-1997; 97US-0049712P.
XX 13-JUN-1997; 97US-0049713P.
XX 13-JUN-1997; 97US-0049714P.
XX 13-JUN-1997; 97US-0049715P.
XX 13-JUN-1997; 97US-0049716P.
XX 13-JUN-1997; 97US-0049717P.
XX 13-JUN-1997; 97US-0049718P.
XX 13-JUN-1997; 97US-0049719P.
XX 13-JUN-1997; 97US-0049720P.
XX 13-JUN-1997; 97US-0049721P.
XX 13-JUN-1997; 97US-0049722P.
XX 13-JUN-1997; 97US-0049723P.
XX 13-JUN-1997; 97US-0049724P.
XX 13-JUN-1997; 97US-0049725P.
XX 13-JUN-1997; 97US-0049726P.
XX 13-JUN-1997; 97US-0049727P.
XX 13-JUN-1997; 97US-0049728P.
XX 13-JUN-1997; 97US-0049729P.
XX 13-JUN-1997; 97US-0049730P.
XX 13-JUN-1997; 97US-0049731P.
XX 13-JUN-1997; 97US-0049732P.
XX 13-JUN-1997; 97US-0049733P.
XX 13-JUN-1997; 97US-0049734P.
XX 13-JUN-1997; 97US-0049735P.
XX 13-JUN-1997; 97US-0049736P.
XX 13-JUN-1997; 97US-0049737P.
XX 13-JUN-1997; 97US-0049738P.
XX 13-JUN-1997; 97US-0049739P.
XX 13-JUN-1997; 97US-0049740P.
XX 13-JUN-1997; 97US-0049741P.
XX 13-JUN-1997; 97US-0049742P.
XX 13-JUN-1997; 97US-0049743P.
XX 13-JUN-1997; 97US-0049744P.
XX 13-JUN-1997; 97US-0049745P.
XX 13-JUN-1997; 97US-0049746P.
XX 13-JUN-1997; 97US-0049747P.
XX 13-JUN-1997; 97US-0049748P.
XX 13-JUN-1997; 97US-0049749P.
XX 13-JUN-1997; 97US-0049750P.
XX 13-JUN-1997; 97US-0049751P.
XX 13-JUN-1997; 97US-0049752P.
XX 13-JUN-1997; 97US-0049753P.
XX 13-JUN-1997; 97US-0049754P.
XX 13-JUN-1997; 97US-0049755P.
XX 13-JUN-1997; 97US-0049756P.
XX 13-JUN-1997; 97US-0049757P.
XX 13-JUN-1997; 97US-0049758P.
XX 13-JUN-1997; 97US-0049759P.
XX 13-JUN-1997; 97US-0049760P.
XX 13-JUN-1997; 97US-0049761P.
XX 13-JUN-1997; 97US-0049762P.
XX 13-JUN-1997; 97US-0049763P.
XX 13-JUN-1997; 97US-0049764P.
XX 13-JUN-1997; 97US-0049765P.
XX 13-JUN-1997; 97US-0049766P.
XX 13-JUN-1997; 97US-0049767P.
XX 13-JUN-1997; 97US-0049768P.
XX 13-JUN-1997; 97US-0049769P.
XX 13-JUN-1997; 97US-0049770P.
XX 13-JUN-1997; 97US-0049771P.
XX 13-JUN-1997; 97US-0049772P.
XX 13-JUN-1997; 97US-0049773P.
XX 13-JUN-1997; 97US-0049774P.
XX 13-JUN-1997; 97US-0049775P.
XX 13-JUN-1997; 97US-0049776P.
XX 13-JUN-1997; 97US-0049777P.
XX 13-JUN-1997; 97US-0049778P.
XX 13-JUN-1997; 97US-0049779P.
XX 13-JUN-1997; 97US-0049780P.
XX 13-JUN-1997; 97US-0049781P.
XX 13-JUN-1997; 97US-0049782P.
XX 13-JUN-1997; 97US-0049783P.
XX 13-JUN-1997; 97US-0049784P.
XX 13-JUN-1997; 97US-0049785P.
XX 13-JUN-1997; 97US-0049786P.
XX 13-JUN-1997; 97US-0049787P.
XX 13-JUN-1997; 97US-0049788P.
XX 13-JUN-1997; 97US-0049789P.
XX 13-JUN-1997; 97US-0049790P.
XX 13-JUN-1997; 97US-0049791P.
XX 13-JUN-1997; 97US-0049792P.
XX 13-JUN-1997; 97US-0049793P.
XX 13-JUN-1997; 97US-0049794P.
XX 13-JUN-1997; 97US-0049795P.
XX 13-JUN-1997; 97US-0049796P.
XX 13-JUN-1997; 97US-0049797P.
XX 13-JUN-1997; 97US-0049798P.
XX 13-JUN-1997; 97US-0049799P.
XX 13-JUN-1997; 97US-0049800P.
XX 13-JUN-1997; 97US-0049801P.
XX 13-JUN-1997; 97US-0049802P.
XX 13-JUN-1997; 97US-0049803P.
XX 13-JUN-1997; 97US-0049804P.
XX 13-JUN-1997; 97US-0049805P.
XX 13-JUN-1997; 97US-0049806P.
XX 13-JUN-1997; 97US-0049807P.
XX 13-JUN-1997; 97US-0049808P.
XX 13-JUN-1997; 97US-0049809P.
XX 13-JUN-1997; 97US-0049810P.
XX 13-JUN-1997; 97US-0049811P.
XX 13-JUN-1997; 97US-0049812P.
XX 13-JUN-1997; 97US-0049813P.
XX 13-JUN-1997; 97US-0049814P.
XX 13-JUN-1997; 97US-0049815P.
XX 13-JUN-1997; 97US-0049816P.
XX 13-JUN-1997; 97US-0049817P.
XX 13-JUN-1997; 97US-0049818P.
XX 13-JUN-1997; 97US-0049819P.
XX 13-JUN-1997; 97US-0049820P.
XX 13-JUN-1997; 97US-0049821P.
XX 13-JUN-1997; 97US-0049822P.
XX 13-JUN-1997; 97US-0049823P.
XX 13-JUN-1997; 97US-0049824P.
XX 13-JUN-1997; 97US-0049825P.
XX 13-JUN-1997; 97US-0049826P.
XX 13-JUN-1997; 97US-0049827P.
XX 13-JUN-1997; 97US-0049828P.
XX 13-JUN-1997; 97US-0049829P.
XX 13-JUN-1997; 97US-0049830P.
XX 13-JUN-1997; 97US-0049831P.
XX 13-JUN-1997; 97US-0049832P.
XX 13-JUN-1997; 97US-0049833P.
XX 13-JUN-1997; 97US-0049834P.
XX 13-JUN-1997; 97US-0049835P.
XX 13-JUN-1997; 97US-0049836P.
XX 13-JUN-1997; 97US-0049837P.
XX 13-JUN-1997; 97US-0049838P.
XX 13-JUN-1997; 97US-0049839P.
XX 13-JUN-1997; 97US-0049840P.
XX 13-JUN-1997; 97US-0049841P.
XX 13-JUN-1997; 97US-0049842P.
XX 13-JUN-1997; 97US-0049843P.
XX 13-JUN-1997; 97US-0049844P.
XX 13-JUN-1997; 97US-0049845P.
XX 13-JUN-1997; 97US-0049846P.
XX 13-JUN-1997; 97US-0049847P.
XX 13-JUN-1997; 97US-0049848P.
XX 13-JUN-1997; 97US-0049849P.
XX 13-JUN-1997; 97US-0049850P.
XX 13-JUN-1997; 97US-0049851P.
XX 13-JUN-1997; 97US-0049852P.
XX 13-JUN-1997; 97US-0049853P.
XX 13-JUN-1997; 97US-0049854P.
XX 13-JUN-1997; 97US-0049855P.
XX 13-JUN-1997; 97US-0049856P.
XX 13-JUN-1997; 97US-0049857P.
XX 13-JUN-1997; 97US-0049858P.
XX 13-JUN-1997; 97US-0049859P.
XX 13-JUN-1997; 97US-0049860P.
XX 13-JUN-1997; 97US-0049861P.
XX 13-JUN-1997; 97US-0049862P.
XX 13-JUN-1997; 97US-0049863P.
XX 13-JUN-1997; 97US-0049864P.
XX 13-JUN-1997; 97US-0049865P.
XX 13-JUN-1997; 97US-0049866P.
XX 13-JUN-1997; 97US-0049867P.
XX 13-JUN-1997; 97US-0049868P.
XX 13-JUN-1997; 97US-0049869P.
XX 13-JUN-1997; 97US-0049870P.
XX 13-JUN-1997; 97US-0049871P.
XX 13-JUN-1997; 97US-0049872P.
XX 13-JUN-1997; 97US-0049873P.
XX 13-JUN-1997; 97US-0049874P.
XX 13-JUN-1997; 97US-0049875P.
XX 13-JUN-1997; 97US-0049876P.
XX 13-JUN-1997; 97US-0049877P.
XX 13-JUN-1997; 97US-0049878P.
XX 13-JUN-1997; 97US-0049879P.
XX 13-JUN-1997; 97US-0049880P.
XX 13-JUN-1997; 97US-0049881P.
XX 13-JUN-1997; 97US-0049882P.
XX 13-JUN-1997; 97US-0049883P.
XX 13-JUN-1997; 97US-0049884P.
XX 13-JUN-1997; 97US-0049885P.
XX 13-JUN-1997; 97US-0049886P.
XX 13-JUN-1997; 97US-0049887P.
XX 13-JUN-1997; 97US-0049888P.
XX 13-JUN-1997; 97US-0049889P.
XX 13-JUN-1997; 97US-0049890P.
XX 13-JUN-1997; 97US-0049891P.
XX 13-JUN-1997; 97US-0049892P.
XX 13-JUN-1997; 97US-0049893P.
XX 13-JUN-1997; 97US-0049894P.
XX 13-JUN-1997; 97US-0049895P.
XX 13-JUN-1997; 97US-0049896P.
XX 13-JUN-1997; 97US-0049897P.
XX 13-JUN-1997; 97US-0049898P.
XX 13-JUN-1997; 97US-0049899P.
XX 13-JUN-1997; 97US-0049900P.
XX 13-JUN-1997; 97US-0049901P.
XX 13-JUN-1997; 97US-0049902P.
XX 13-JUN-1997; 97US-0049903P.
XX 13-JUN-1997; 97US-0049904P.
XX 13-JUN-1997; 97US-0049905P.
XX 13-JUN-1997; 97US-0049906P.
XX 13-JUN-1997; 97US-0049907P.
XX 13-JUN-1997; 97US-0049908P.
XX 13-JUN-1997; 97US-0049909P.
XX 13-JUN-1997; 97US-0049910P.
XX 13-JUN-1997; 97US-0049911P.
XX 13-JUN-1997; 97US-0049912P.
XX 13-JUN-1997; 97US-0049913P.
XX 13-JUN-1997; 97US-0049914P.
XX 13-JUN-1997; 97US-0049915P.
XX 13-JUN-1997; 97US-0049916P.
XX 13-JUN-1997; 97US-0049917P.
XX 13-JUN-1997; 97US-0049918P.
XX 13-JUN-1997; 97US-0049919P.
XX 13-JUN-1997; 97US-0049920P.
XX 13-JUN-1997; 97US-0049921P.
XX 13-JUN-1997; 97US-0049922P.
XX 13-JUN-1997; 97US-0049923P.
XX 13-JUN-1997; 97US-0049924P.
XX 13-JUN-1997; 97US-0049925P.
XX 13-JUN-1997; 97US-0049926P.
XX 13-JUN-1997; 97US-0049927P.
XX 13-JUN-1997; 97US-0049928P.
XX 13-JUN-1997; 97US-0049929P.
XX 13-JUN-1997; 97US-0049930P.
XX 13-JUN-1997; 97US-0049931P.
XX 13-JUN-1997; 97US-0049932P.
XX 13-JUN-1997; 97US-0049933P.
XX 13-JUN-1997; 97US-0049934P.
XX 13-JUN-1997; 97US-0049935P.
XX 13-JUN-1997; 97US-0049936P.
XX 13-JUN-1997; 97US-0049937P.
XX 13-JUN-1997; 97US-0049938P.
XX 13-JUN-1997; 97US-0049939P.
XX 13-JUN-1997; 97US-0049940P.
XX 13-JUN-1997; 97US-0049941P.
XX 13-JUN-1997; 97US-0049942P.
XX 13-JUN-1997; 97US-0049943P.
XX

Db 10 NKSESDKXVFLVYKTLQAQLQVNTMLEYLIDVEIARSDCKRPLSTN 55

RESULT 14

ID ADA57563 standard; protein; 113 AA.

ADA57563;

20-NOV-2003 (first entry)

Human secreted protein #514.

immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
cytotoxic; cerebroprotective; neuroprotective; nootropic;
cardiovascular; antiarteriosclerotic; gene therapy;
human secreted protein; immune disorder; inflammation;
respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
multiple sclerosis; ischaemic brain injury; Parkinson's disease;
Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
triple helix formation; antisense gene therapy; forensic biology.

Homo sapiens.

WO2002102994-A2.

27-DEC-2002.

19-MAR-2002; 2002WO-US008278.

21-MAR-2001; 2001US-0277340P.

19-JUL-2001; 2001US-0306171P.

13-NOV-2001; 2001US-0331287P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI, 2003-167512/16.

N-PSDB; ADA56670.

New human secreted polypeptides and polynucleotides, useful for
diagnosing, treating or preventing e.g. immune disorders, inflammatory
conditions, respiratory disorders, cancers, CNS disorders, or
neurodegenerative disorders.

Claim 13; SEQ ID NO 1756; 1754bp; English.

The invention relates to 592 new human secreted polypeptides useful for
diagnosing, treating or preventing e.g. immune disorders, inflammatory
conditions, respiratory disorders, cancers, CNS disorders, or
neurodegenerative disorders, or polypeptides comprising an amino acid
sequence at least 95% identical to the new sequences. The polypeptides,
antibodies or antibody fragments that bind to the polypeptides, nucleic
acids encoding the polypeptides, agonists or antagonists that binds to
the polypeptide, are useful in preparing diagnostic or pharmaceutical
compositions for diagnosing, treating or preventing an e.g. immune
disorder, inflammatory conditions (e.g. inflammatory bowel disease,
nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
(e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
disorders (e.g. Parkinson's disease or Alzheimer's disease), and
cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
polynucleotides are useful for chromosome identification, chromosome
mapping, for controlling gene expression through triple helix formation
or antisense DNA or RNA, in gene therapy, for identifying individuals
from minute biological samples, in forensic biology, and as hybridization
probes. The polypeptides are useful for as molecular weight markers on
sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
gels, to raise antibodies, for testing biological activities, and for
treating or preventing neural disorders, immune system disorders,

muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
renal, proliferative and/or cancerous diseases. This sequence corresponds
to one of the polypeptide of the invention. Note: The sequence data for
this patent did form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Sequence 113 AA;

Query Match 46.5%; Score 118; DB 6; Length 113;
Best Local Similarity 50.0%; Pred. No. 1,1e-08;
Matches 23; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

1 NKSESDKXHFRIFRVLKVRQVTDHLEHYHNVEMQWTTCKQPERTN 46

58 NKSESDKXVFLVYKTLQAQLQVNTMLEYLIDVEIARSDCKRPLSTN 103

RESULT 15

ID ADA41457 standard; protein; 113 AA.

ADA41457;

20-NOV-2003 (first entry)

Human secreted protein.

Human, secreted protein; cancer; hyperproliferative disorder;
rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
anaemia; allergic reaction; asthma; cardiovascular disorder;
wound healing; cytostatic; immunosuppressive; nootropic; neuroprotective;
antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
vulnerary; cardiac; gene therapy.

Homo sapiens.

WO2002102993-A2.

27-DEC-2002.

19-MAR-2002; 2002WO-US008123.

21-MAR-2001; 2001US-0277340P.

19-JUL-2001; 2001US-0306171P.

13-NOV-2001; 2001US-0331287P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI, 2003-175238/17.

New human secreted proteins and nucleic acid molecules, useful for
preparing a diagnostic or pharmaceutical composition for diagnosing,
preventing or treating cancer or other hyperproliferative disorder,
asthma, allergies or AIDS.

Claim 1; SEQ ID NO 1840; 3205bp; English.

The invention relates to novel genes ADA39629-ADA40565 and proteins
ADA40566-ADA41501 for human secreted proteins, useful for preventing,
treating or ameliorating medical conditions e.g. by protein or gene
therapy. The polypeptides, nucleic acid molecules, antibodies or their
fragments, and agonists or antagonists that bind to the polypeptide are
useful for preparing a diagnostic or pharmaceutical composition for
diagnosing or treating cancer or other hyperproliferative disorder. The
polypeptides and nucleic acid molecules are also useful for detecting,
preventing, diagnosing, prognosticating, treating or ameliorating cancer
or other hyperproliferative disorders including neoplasms, autoimmune
disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
anaemia), haematopoietic or haematological disorders (e.g. anaemia,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 11.7738 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254
Sequence: 1 NKESSDKYHFRIRVLKQVR.....EYHLNVMQWTCCKPETTN 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCUS COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	46.5	142	3	US-09-431-480-4
2	118	46.5	142	3	US-09-617-302-4
3	113	44.5	141	3	US-09-431-480-2
4	113	44.5	141	3	US-09-617-302-2
5	111	43.7	142	3	US-09-431-480-3
6	111	43.7	142	3	US-09-617-302-3
7	103	40.6	127	4	US-08-849-303-19
8	99	39.0	140	4	US-09-886-319A-46
9	99	39.0	140	4	US-09-886-319A-48
10	98	38.6	112	4	US-09-886-303-16
11	98	38.6	118	4	US-09-775-932-24
12	95	37.4	140	3	US-09-431-480-5
13	95	37.4	140	3	US-09-617-302-5
14	95	37.4	140	4	US-08-849-303-18
15	94	37.0	116	4	US-09-775-932-16
16	94	37.0	120	4	US-09-775-932-2
17	94	37.0	120	6	US-09-775-932-2
18	94	37.0	139	2	US-08-791-522-4
19	94	37.0	139	3	US-09-314-777-4
20	94	37.0	139	4	US-08-849-303-15
21	94	37.0	145	2	US-08-832-535-11
22	94	37.0	146	2	US-08-791-522-3
23	94	37.0	146	3	US-08-744-138-3
24	94	37.0	146	3	US-09-019-485-4
25	94	37.0	146	3	US-09-314-777-3
26	94	37.0	146	3	US-09-431-480-6
27	94	37.0	146	3	US-09-617-302-6

28	94	37.0	146	4	US-09-241-376-3	Sequence 3, Appl1
29	94	37.0	146	4	US-09-528-436B-3	Sequence 3, Appl1
30	94	37.0	146	4	US-09-886-319A-47	Sequence 47, Appl1
31	94	37.0	146	4	US-09-940-497-3	Sequence 3, Appl1
32	94	37.0	146	4	US-09-976-594-37	Sequence 37, Appl1
33	94	37.0	146	4	US-08-849-303-17	Sequence 17, Appl1
34	94	37.0	146	5	PCT-US95-07135-9	Sequence 9, Appl1
35	94	37.0	146	6	US-09-775-932-2	Patent No. 5432264
36	84	33.1	121	4	US-09-775-932-8	Sequence 8, Appl1
37	84	33.1	111	3	US-08-744-138-6	Sequence 6, Appl1
38	84	33.1	141	4	US-09-241-376-6	Sequence 6, Appl1
39	84	33.1	141	4	US-09-940-497-6	Sequence 24, Appl1
40	84	33.1	141	4	US-08-849-303-24	Sequence 24, Appl1
41	81	31.9	162	4	US-08-849-303-25	Sequence 25, Appl1
42	80	31.5	111	4	US-08-849-303-26	Sequence 26, Appl1
43	78	30.7	121	4	US-09-775-932-14	Sequence 14, Appl1
44	78	30.7	128	4	US-09-775-932-12	Sequence 12, Appl1
45	78	30.7	149	2	US-08-461-030C-2	Sequence 2, Appl1

ALIGNMENTS

```

RESULT 1
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ. ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match      46.5%; Score 118; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 4.5e-10;
Matches 23; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

CY      1 NKESSDKYHFRIRVLKQVRQVTDHLEYHLNVMQWTCCKPETTN 46
Db      57 NKESSDKYHFRIRVLKQVRQVTDHLEYHLNVMQWTCCKPETTN 102

RESULT 2
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 142
TYPE: PRT
ORGANISM: Homo sapiens
US-09-617-302-4

Query Match 46.5%; Score 118; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 4.5e-10;
Matches 23; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVGQVTDHLEHNLVEMQWTTCKP 46
DB 57 NKESDXYVFLVVKTLQALQVNLLEYLIDVELIARSDCRKPLSTN 102

RESULT 3
US-09-431-480-2
Sequence 2, Application US/09431480
Patent No. 6235708
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72
CURRENT APPLICATION NUMBER: US/09/431,480
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 60/109,217
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/156,382
EARLIER FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-431-480-2

Query Match 44.5%; Score 113; DB 3; Length 141;
Best Local Similarity 51.2%; Pred. No. 2.5e-09;
Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVGQVTDHLEHNLVEMQWTTCKP 41
DB 56 NKASNDLYNFRVVDILKSQEQITDSLLEYLVNIARTWCKK 96

RESULT 4
US-09-617-302-2
Sequence 2, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-617-302-2

Query Match 44.5%; Score 113; DB 3; Length 141;
Best Local Similarity 51.2%; Pred. No. 2.5e-09;
Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVGQVTDHLEHNLVEMQWTTCKP 41
DB 56 NKASNDLYNFRVVDILKSQEQITDSLLEYLVNIARTWCKK 96

RESULT 5
US-09-431-480-3
Sequence 3, Application US/09431480
Patent No. 6235708
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72
CURRENT APPLICATION NUMBER: US/09/431,480
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 60/109,217
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/156,382
EARLIER FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 142
TYPE: PRT
ORGANISM: Mus musculus
US-09-431-480-3

Query Match 43.7%; Score 111; DB 3; Length 142;
Best Local Similarity 42.9%; Pred. No. 5.1e-09;
Matches 18; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVGQVTDHLEHNLVEMQWTTCKP 42
DB 57 NKESDXYVFLVVKTLQALQVNLLEYLIDVELIARSDCRKPLSTN 102

RESULT 6
US-09-617-302-3
Sequence 3, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 142
TYPE: PRT
ORGANISM: Mus musculus
US-09-617-302-3

Query Match 43.7%; Score 111; DB 3; Length 142;
Best Local Similarity 42.9%; Pred. No. 5.1e-09;
Matches 18; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVGQVTDHLEHNLVEMQWTTCKP 42
DB 57 NKESDXYVFLVVKTLQALQVNLLEYLIDVELIARSDCRKPLSTN 102

RESULT 7

US-08-849-303-19
 ; Sequence 19, Application US/08849303
 ; Patent No. 6680424
 ; GENERAL INFORMATION:
 ; APPLICANT: Atkinson, Howard J.
 ; APPLICANT: McPherson, Michael J.
 ; APPLICANT: Urwin, Peter E.
 ; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
 ; NUMBER OF SEQUENCES: 79
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue, 4th Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/849,303
 ; FILING DATE: 21-MAY-1997
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 1321-1-003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 127 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; US-08-849-303-19

Query Match 40.6%; Score 103; DB 4; Length 127;
 Best Local Similarity 43.2%; Pred. No. 7,1e-08;
 Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLTQVROVTDHLEHYNEMQWTTCKPPT 44
 Db 42 NKSNDAYHSRAIQVVRARQKLVAGVNYFLDVEWGRTTCKSQ 85

RESULT 8

US-09-886-319A-46
 ; Sequence 46, Application US/09886319A
 ; Patent No. 6586185
 ; GENERAL INFORMATION:
 ; APPLICANT: Wolf, Eckard
 ; APPLICANT: Werner, Sabine
 ; APPLICANT: Haile, Jörn-Peter
 ; APPLICANT: Regenbogen, Johannes
 ; APPLICANT: Goppelt, Andreas
 ; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
 ; TITLE OF INVENTION: the diagnosis or treatment of skin disorders and wound
 ; TITLE OF INVENTION: Healing and for the identification of pharmacologically
 ; TITLE OF INVENTION: Active Substances
 ; FILE REFERENCE: 50125/014002
 ; CURRENT APPLICATION NUMBER: US/09/886,319A
 ; CURRENT FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: US 60/222,081

;; PRIOR FILING DATE: 2000-08-01
 ; PRIOR APPLICATION NUMBER: DE 10030149.5
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 140
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-886-319A-46

Query Match 39.0%; Score 99; DB 4; Length 140;
 Best Local Similarity 43.2%; Pred. No. 3,2e-07;
 Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLTQVROVTDHLEHYNEMQWTTCKPPT 44
 Db 55 NKSNDAYHSRAIQVVRARQKLVAGVNYFLDVEWGRTTCKSQ 98

RESULT 9

US-09-886-319A-48
 ; Sequence 48, Application US/09886319A
 ; Patent No. 6586185
 ; GENERAL INFORMATION:
 ; APPLICANT: Wolf, Eckard
 ; APPLICANT: Werner, Sabine
 ; APPLICANT: Haile, Jörn-Peter
 ; APPLICANT: Regenbogen, Johannes
 ; APPLICANT: Goppelt, Andreas
 ; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
 ; TITLE OF INVENTION: the diagnosis or treatment of skin disorders and wound
 ; TITLE OF INVENTION: Healing and for the identification of pharmacologically
 ; TITLE OF INVENTION: Active Substances
 ; FILE REFERENCE: 50125/014002
 ; CURRENT APPLICATION NUMBER: US/09/886,319A
 ; CURRENT FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: US 60/222,081
 ; PRIOR FILING DATE: 2000-08-01
 ; PRIOR APPLICATION NUMBER: DE 10030149.5
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 48
 ; LENGTH: 140
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-886-319A-48

Query Match 39.0%; Score 99; DB 4; Length 140;
 Best Local Similarity 43.2%; Pred. No. 3,2e-07;
 Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKSDDKTHFRIRVLTQVROVTDHLEHYNEMQWTTCKPPT 44
 Db 55 NKSNDAYHSRAIQVVRARQKLVAGVNYFLDVEWGRTTCKSQ 98

RESULT 10

US-08-849-303-16
 ; Sequence 16, Application US/08849303
 ; Patent No. 6680424
 ; GENERAL INFORMATION:
 ; APPLICANT: Atkinson, Howard J.
 ; APPLICANT: McPherson, Michael J.
 ; APPLICANT: Urwin, Peter E.
 ; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
 ; NUMBER OF SEQUENCES: 79
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue, 4th Floor
 ; CITY: Hackensack
 ; STATE: New Jersey

```

; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-08-849-303-16

```

```

Query Match          38.6%; Score 98; DB 4; Length 112;
Best Local Similarity 41.9%; Pred. No. 3.4e-07;
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKXHFRIPIVLYKQROVTDHLEHNLNVMQMTTCCKPE 43
DB      28 NKSNDAYQSRVAVVAVARAKQVVGMYFLDVELGRITCTKSQ 70

```

```

RESULT 11
US-09-775-932-24
; Sequence 24, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cycatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 24
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Bos taurus
; US-09-775-932-24

```

```

Query Match          38.6%; Score 98; DB 4; Length 118;
Best Local Similarity 41.9%; Pred. No. 3.6e-07;
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKXHFRIPIVLYKQROVTDHLEHNLNVMQMTTCCKPE 43
DB      34 NKSNDAYQSRVAVVAVARAKQVVGMYFLDVELGRITCTKSQ 76

```

```

RESULT 12
US-09-431-480-5
; Sequence 5, Application US/09431480
; Patent No. 6235708

```

```

; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-431-480-5

```

```

Query Match          37.4%; Score 95; DB 3; Length 140;
Best Local Similarity 40.9%; Pred. No. 1.3e-06;
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKXHFRIPIVLYKQROVTDHLEHNLNVMQMTTCCKPE 44
DB      55 NKSNDAYHSRAIQVAVARAKQVAVGVNFFDVEGRITCTKSQT 98

```

```

RESULT 13
US-09-617-302-5
; Sequence 5, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-617-302-5

```

```

Query Match          37.4%; Score 95; DB 3; Length 140;
Best Local Similarity 40.9%; Pred. No. 1.3e-06;
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKXHFRIPIVLYKQROVTDHLEHNLNVMQMTTCCKPE 44
DB      55 NKSNDAYHSRAIQVAVARAKQVAVGVNFFDVEGRITCTKSQT 98

```

```

RESULT 14
US-08-849-303-18
; Sequence 18, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urvinn, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:

```

Search completed: March 18, 2004, 14:25:36
Job time : 11.7738 secs

ADDRESSEE: Klaubert & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-18

Query Match 37.4%; Score 95; DB 4; Length 140;
Best Local Similarity 40.9%; Pred. No. 1.3e-06;
Matches 18; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIPRVLVKQROVTDHLEHNLVEMQWTTCKP 44
DB 55 NKGSNDAYHSRAIQVRRKQLVAGVNYFFDVENGRITTCR 98

RESULT 15
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cytatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match 37.0%; Score 94; DB 4; Length 116;
Best Local Similarity 43.9%; Pred. No. 1.4e-06;
Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIPRVLVKQROVTDHLEHNLVEMQWTTCK 41
DB 33 NRASNDKYSRVRVVISAKROLVSGIKYIIQVEIGRTTCRK 73

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 27.8373 Seconds
(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254
Sequence: 1 NKESDQKXHFRIPLVKVOR.....EYHLNVQMWTCKPPTTN 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCFUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	46	9	US-09-941-314-10 Sequence 10, Appl
2	254	100.0	49	9	US-09-941-314-13 Sequence 13, Appl
3	254	100.0	80	9	US-09-941-314-15 Sequence 15, Appl
4	254	100.0	115	9	US-09-941-314-3 Sequence 3, Appl
5	254	100.0	117	9	US-09-941-314-4 Sequence 4, Appl
6	254	100.0	137	9	US-09-941-314-2 Sequence 2, Appl
7	248	97.6	52	9	US-09-941-314-14 Sequence 14, Appl
8	154	60.6	59	9	US-09-941-314-16 Sequence 16, Appl
9	149	58.7	33	9	US-09-941-314-12 Sequence 12, Appl
10	138	54.3	24	9	US-09-941-314-11 Sequence 11, Appl
11	125	49.2	27	9	US-09-941-314-9 Sequence 9, Appl
12	125	49.2	35	9	US-09-941-314-8 Sequence 8, Appl
13	110	43.3	50	9	US-09-864-761-34822 Sequence 34822, A
14	110	43.3	50	9	US-09-864-761-48936 Sequence 48936, A
15	107	42.1	145	9	US-09-740-638-2 Sequence 2, Appl

16	107	42.1	145	13	US-10-006-467-2 Sequence 2, Appl
17	107	42.1	145	14	US-10-235-148-2 Sequence 2, Appl
18	103	40.6	127	8	US-08-849-303-19 Sequence 19, Appl
19	101	39.8	145	14	US-10-168-425-14 Sequence 14, Appl
20	99	39.0	140	14	US-10-376-564-46 Sequence 46, Appl
21	99	39.0	140	14	US-10-376-564-48 Sequence 48, Appl
22	98	38.6	112	8	US-08-849-303-16 Sequence 16, Appl
23	98	38.6	118	9	US-09-775-932-24 Sequence 24, Appl
24	98	38.6	101	15	US-10-264-046-2608 Sequence 2608, Ap
25	95	37.4	140	8	US-08-849-303-18 Sequence 18, Appl
26	94	37.0	116	9	US-09-775-932-16 Sequence 16, Appl
27	94	37.0	120	9	US-09-775-932-2 Sequence 2, Appl
28	94	37.0	139	8	US-08-849-303-15 Sequence 15, Appl
29	94	37.0	139	9	US-09-969-834-4 Sequence 4, Appl
30	94	37.0	146	8	US-08-849-303-17 Sequence 17, Appl
31	94	37.0	146	9	US-09-940-497-3 Sequence 3, Appl
32	94	37.0	146	9	US-09-969-834-3 Sequence 3, Appl
33	94	37.0	146	14	US-10-329-428-3 Sequence 3, Appl
34	94	37.0	146	14	US-10-376-564-47 Sequence 47, Appl
35	93	36.6	48	9	US-09-941-314-17 Sequence 17, Appl
36	87	34.3	165	9	US-09-740-638-5 Sequence 5, Appl
37	87	34.3	165	13	US-10-006-467-5 Sequence 5, Appl
38	87	34.3	165	14	US-10-235-148-5 Sequence 5, Appl
39	84	33.1	121	9	US-09-775-932-8 Sequence 8, Appl
40	84	33.1	141	8	US-08-849-303-24 Sequence 24, Appl
41	84	33.1	141	9	US-09-940-497-6 Sequence 6, Appl
42	81	31.9	162	8	US-08-849-303-25 Sequence 25, Appl
43	80	31.5	111	8	US-08-849-303-26 Sequence 26, Appl
44	78	30.7	121	9	US-09-775-932-14 Sequence 14, Appl
45	78	30.7	128	9	US-09-775-932-12 Sequence 12, Appl

ALIGNMENTS

```

RESULT 1
US-09-941-314-10
; Sequence 10, Application US/09941314
; Patient No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941, 314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230, 230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-10

Query Match      100.0%; Score 254; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2, 2e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 NKESDQKXHFRIPLVKVORQVTDHLEHYHLNVQMWTCKPPTTN 46
Db      1 NKESDQKXHFRIPLVKVORQVTDHLEHYHLNVQMWTCKPPTTN 46

RESULT 2
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patient No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941, 314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230, 230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-10

Query Match      100.0%; Score 254; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2, 2e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 49
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-13

Query Match 100.0%; Score 254; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46
DB 4 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 49

RESULT 3
US-09-941-314-15
Sequence 15, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 80
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-15

Query Match 100.0%; Score 254; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.2e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46
DB 3 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 48

RESULT 4
US-09-941-314-3
Sequence 3, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-3

Query Match 100.0%; Score 254; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 6.4e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46
DB 33 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 78

RESULT 5
US-09-941-314-4
Sequence 4, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-4

Query Match 100.0%; Score 254; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.5e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46
DB 35 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 80

RESULT 6
US-09-941-314-2
Sequence 2, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 137
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-2

Query Match 100.0%; Score 254; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 7.7e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 46
DB 55 NKESDDKYHFRIRFVLKVQRQVTDHLEHYHNVEMQWTTCKPPTN 100

RESULT 7
US-09-941-314-14
Sequence 14, Application US/09941314
Patent No. US20020142396A1

```
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-14

Query Match      97.6%; Score 248; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 1,7e-26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 KESDDKYHFRIRFVLKQROVTDHLEHNLNEMQWTTCKRPETN 46
Db      1 KESDDKYHFRIRFVLKQROVTDHLEHNLNEMQWTTCKRPETN 45

RESULT 8
US-09-941-314-16
; Sequence 16, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-16

Query Match      60.6%; Score 154; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 1,4e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 RQVTDHLEHNLNEMQWTTCKRPETN 46
Db      1 RQVTDHLEHNLNEMQWTTCKRPETN 27

RESULT 9
US-09-941-314-12
; Sequence 12, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 33
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-12

Query Match      58.7%; Score 149; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 3,6e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 QVTDHLEHNLNEMQWTTCKRPETN 46
Db      1 QVTDHLEHNLNEMQWTTCKRPETN 26

RESULT 10
US-09-941-314-11
; Sequence 11, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-11

Query Match      54.3%; Score 138; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 8e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 RQVTDHLEHNLNEMQWTTCKRPE 43
Db      1 RQVTDHLEHNLNEMQWTTCKRPE 24

RESULT 11
US-09-941-314-9
; Sequence 9, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-9

Query Match      49.2%; Score 125; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 5,5e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NKESDDKYHFRIRFVLKQROVTD 24
Db      4 NKESDDKYHFRIRFVLKQROVTD 27
```

```
RESULT 12
US-09-941-314-8
; Sequence 8, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and its use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FaSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-8

Query Match          49.2%; Score 125; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 7.4e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NKSDDKXHFRIFRVLKVRQVTD 24
        ||| ||| ||| ||| ||| |||
Db      12 NKSDDKXHFRIFRVLKVRQVTD 35

RESULT 13
US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34822
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109954.10
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUATE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match          43.3%; Score 110; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NKSDDKXHFRIFRVLKVRQ 21
        ||| ||| ||| ||| ||| |||
Db      30 NKSDDKXHFRIFRVLKVRQ 50

RESULT 14
US-09-864-761-48936
; Sequence 48936, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```


Fri Mar 19 07:21:53 2004

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48936
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096677.18
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: A1300857.1, EVALUE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01
US-09-864-761-48936

```

```

Query Match          43.3%; Score 110; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKHYFRIRPVLYKVQRQ 21
Db      30 NKSDDKHYFRIRPVLYKVQRQ 50

```

```

RESULT 15
US-09-740-638-2
; Sequence 2, Application US/09740638
; Patent No. US20020006656A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Zcy5: A Member of the Cystatin
; FILE REFERENCE: 99-104
; CURRENT APPLICATION NUMBER: US/09/740,638
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-740-638-2

```

```

Query Match          42.1%; Score 107; DB 9; Length 145;
Best Local Similarity 39.1%; Pred. No. 1.1e-06;
Matches 18; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

```

```

QY      1 NKSDDKHYFRIRPVLYKVQRQVTDHLEHNLNEMQWTTCKRPETN 46
Db      53 NNASNDLYLVQVRLIRSORQLTGVGEIVTVKIGWTKCRNDTSN 98

```

Search completed: March 18, 2004, 14:30:54
Job time : 28.8373 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 8.85317 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254

Sequence: 1 NKSSEDKYHFRIRVLKQVR.....EYHLNVEMQWTTQCKPETT 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 78:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	43.7	139	2	A45361
2	103	40.6	120	2	S10587
3	103	40.6	127	2	S07085
4	98	38.6	112	1	UDBO
5	95	37.4	140	2	A36163
6	94	37.0	139	1	UDCH
7	94	37.0	146	1	UDHU
8	84	33.1	141	2	B29632
9	81	31.9	162	2	A43428
10	80	31.5	111	2	A28793
11	75	29.5	133	2	JC4536
12	73	28.7	141	1	UDHDP2
13	71	28.0	111	1	JC2040
14	71	28.0	132	2	CY4918
15	71	28.0	141	1	UDHUP1
16	71	28.0	149	2	T31871
17	67	26.4	142	2	A47142
18	63	24.8	438	2	A47702
19	62	24.4	438	1	T52149
20	62	24.4	434	1	KGBO12
21	62	24.4	439	1	KGBOH2
22	62	24.0	436	1	KGBO11
23	61	24.0	621	1	KGBOH1
24	60	23.6	257	2	T03724
25	59	23.2	139	2	B64005
26	59	23.2	644	1	KGHUI1
27	59	23.2	644	1	KGHUI1
28	58	23.0	246	2	H90539
29	58	22.8	448	2	JN0118

30	57.5	22.6	1585	2	T19121	probable protein-t
31	57	22.4	141	2	JQ1470	cystatin S precurs
32	56.5	22.2	325	2	P69784	conserved hypochet
33	56	22.0	200	2	I46051	secreted phosphopr
34	56	22.0	588	2	C95252	L-fucose isomerase
35	56	22.0	588	2	A97117	L-fucose isomerase
36	56	22.0	588	2	S19254	nitrate reductase
37	55.5	21.9	164	2	T31026	hypothetical prote
38	55	21.7	480	2	H86411	protein PIK23.12 (
39	55	21.7	527	2	E90740	probable enzyme [i
40	55	21.7	527	2	G85590	probable enzyme yb
41	55	21.7	527	2	G64818	probable membrane
42	55	21.7	904	1	RDNTNT	nitrate reductase
43	55	21.7	904	1	RDNTNT	nitrate reductase
44	54.5	21.5	1421	2	T49500	hypothetical prote
45	54	21.3	226	2	F83782	transcription regu

ALIGNMENTS

RESULT 1

A45361 cystatin-related epididymal specific protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A45361

R:Corwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.

Mol. Endocrinol. 6, 1653-1664, 1992

A:Title: The CRIS gene: a unique testis-regulated gene related to the cystatin family is

A:Reference number: A45361; PMID:93078799; PMID:1280328

A:Accession: A45361

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-139 <COR>

A:Cross-references: GB:549926; NID:9260492; PIDN:AAC53390.1; PID:9260493

A:Note: sequence extracted from NCBI backbone (NCBIP:118813)

C:Superfamily: cystatin; cystatin homology

F:28-139/Domain: cystatin homology <CYS>

Query Match 43.7%; Score 111; DB 2; Length 139;

Best Local Similarity 42.9%; Pred. No. 2.7e-07;

Matches 18; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Db 54 NKSSEDKYHFRIRVLKQVRQVTDHLEHLNVEMQWTTQCKP 42

RESULT 2

S10587

cystatin C - rat

C:Species: Rattus sp. (rat)

C>Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999

C:Accession: S10587

R:Esnard, F.; Esnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billiard, M.; Gauch.

Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990

A:Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation

A:Reference number: S10587; PMID:90380276; PMID:2400577

A:Accession: S10587

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-120 <ESN>

A:Note: 43-Asn was also found

A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala

F:9-120/Domain: cystatin homology <CYS>

Query Match 40.6%; Score 103; DB 2; Length 120;

Best Local Similarity 43.2%; Pred. No. 2.7e-06;

Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Db 1 NKSSEDKYHFRIRVLKQVRQVTDHLEHLNVEMQWTTQCKP 44

A:Reference number: S48159; MUID:95010016; PMID:7925354
 A:Accession: S48159
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 24-139 <AB>
 R:Lieber, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.
 FEBS Lett. 249, 162-168, 1989
 A:Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.
 A:Reference number: S04008; MUID:89252033; PMID:2721673
 A:Accession: S04008
 A:Molecule type: protein
 A:Residues: 97-114 <LAB>
 R:Colletta, R.; Bird, J.W.C.
 Gene 130, 175-181, 1993
 A:Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping
 A:Reference number: JN0789; MUID:93366172; PMID:8359684
 A:Accession: JN0789
 A:Molecule type: DNA
 A:Residues: 1-139 <C02>
 A:Cross-references: GB:M95725
 A:Note: authors failed to translate the codon for residue 115-Tyr
 C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases
 C:Genetics:
 A:Gene: Ccn
 A:Introns: 76/3; 114/3
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-139/Product: cystatin, long form #status experimental <CYLF>
 F:30-139/Domain: cystatin homology <CYS>
 F:32-139/Product: cystatin, short form #status experimental <CYSR>
 F:76-80/Region: inhibitory #status predicted
 F:94-104,118-138/Diulfidic bonds: #status experimental
 F:103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 37.0%; Score 94; DB 1; Length 139;
 Best Local Similarity 43.9%; Pred. No. 4.8e-05;
 Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKYPFRIFRVLKVRQVTDHLEHLMVEMQMTTCOK 41
 Db 56 NRASNDKYSRRVVRVIAKQQLVSGIKYILQVEIGRTTCRK 96

RESULT 7
 UDHU
 Cystatin C precursor [validated] - human
 N:Alternate names: gamma-CSP; gamma-trace; neuroendocrine basic polypeptide; post-gamma
 C:Species: Homo sapiens (man)
 C>Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000
 C:Accession: S10216; S00004; J10095; A33400; S02751; A01270; A25434; S12288; A32732; A60
 R:Abrahamson, M.; Olafsson, I.; Palssdottrir, A.; Ulvbaeck, M.; Lundvall, A.; Jansson, O.
 Biochem. J. 268, 287-294, 1990
 A:Title: Structure and expression of the human cystatin C gene.
 A:Reference number: S10216; MUID:90303202; PMID:2363674
 A:Accession: S10216
 A:Molecule type: DNA
 A:Residues: 1-146 <AB1>
 A:Cross-references: EMBL:X52255; NID:930257; PIDN:CAA3497.1; PID:g296643
 R:Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundvall, A.
 FEBS Lett. 216, 229-233, 1987
 A:Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the
 A:Reference number: S00004; MUID:87219149; PMID:3495457
 A:Accession: S00004
 A:Molecule type: mRNA
 A:Residues: 1-146 <AB2>
 A:Cross-references: EMBL:X05607; NID:930371; PIDN:CAA2096.1; PID:g755738
 R:Levy, E.; Lopez-Otin, C.; Ghiso, J.; Geltnier, D.; Frangione, B.
 J. Exp. Med. 169, 1771-1778, 1989
 A:Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a
 A:Reference number: J10095; MUID:89235594; PMID:2541223
 A:Accession: J10095
 A:Molecule type: DNA

A:Residues: 1-146 <LEV>
 A:Cross-references: GB:X61681; NID:930367; PIDN:CAA43856.2; PID:g4490944
 A:Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit
 e)
 R:Satch, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, H.
 Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989
 A:Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which
 A:Reference number: A33400; MUID:89350949; PMID:2764935.
 A:Accession: A33400
 A:Molecule type: DNA
 A:Residues: 1-24, 'T', 26-146 <SNI>
 A:Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:9181385; PIDN:AA52164.1; PID:g91
 R:Ghisso, J.; Cowan, N.; Frangione, B.
 Biol. Chem. Hoppe-Seyler 369, 205-208, 1988
 A:Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron
 A:Reference number: S02751; MUID:89076507; PMID:3264504
 A:Accession: S02751
 A:Molecule type: DNA
 A:Residues: 82-119 <GH2>
 A:Cross-references: EMBL:M2769
 A:Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown
 R:Grubb, A.; Lofberg, H.
 Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982
 A:Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the
 A:Reference number: A01270; MUID:82222268; PMID:6283552
 A:Accession: A01270
 A:Molecule type: protein
 A:Residues: 27-111, 'S', 133-146 <GRU>
 R:Ghisso, J.; Jansson, O.; Frangione, B.
 Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986
 A:Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland ty
 A:Reference number: A25434; MUID:86206076; PMID:3517880
 A:Accession: A25434
 A:Molecule type: protein
 A:Residues: 37-93, 'Q', 95-146 <GHI>
 R:Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Brodtkin, M.; Borchardt, U.; Machleidt, W.
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983
 A:Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatir
 A:Reference number: S01461; MUID:84110053; PMID:6662458
 A:Accession: S12288
 A:Molecule type: protein
 A:Residues: 27-73 <TUR>
 R:Brzin, J.; Popovic, T.; Turk, V.
 Biochem. Biophys. Res. Commun. 118, 103-109, 1984
 A:Title: Human cystatin, a new protein inhibitor of cysteine proteinases.
 A:Reference number: A32732; MUID:84128015; PMID:6365094
 A:Accession: A32732
 A:Molecule type: protein
 A:Residues: 27-76 <BR2>
 R:Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.
 Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
 A:Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst
 A:Reference number: A60552; MUID:90193615; PMID:2315647
 A:Accession: A60552
 A:Molecule type: protein
 A:Residues: 27-49, 'X', 52-64 <OLA>
 A:Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
 e defective gene is not present in CSF but is found instead in amyloid deposits
 R:Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
 Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
 A:Title: Different forms of human cystatin C.
 A:Reference number: S10607; MUID:91025625; PMID:2222856
 A:Accession: S10607
 A:Molecule type: protein
 A:Residues: 27-53 <POP>
 A:Experimental source: urine, kidney disease
 A:Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also
 R:Grubb, A.; Lofberg, H.; Barrett, A.J.
 FEBS Lett. 170, 370-374, 1984
 A:Title: The disulfide bridges of human cystatin C (gamma-trace) and chicken cystatin.
 A:Reference number: S01462
 A:Contents: annotation; disulfide bonds
 R:Berti, P.J.; Storer, A.C.

Biochem. J. 302, 411-416, 1994
 A/Title: Local pH-dependent conformational changes leading to proteolytic susceptibility
 A/Reference number: S55305; MUID:94379969; PMID:8092991
 A/Accession: S55305
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 27-49/106-146 <BER>
 C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fluid patients with certain autoimmune diseases.
 C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an important role in the regulation of the immune response.
 C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorrhage with amyloidosis (Finnish type).
 C/Genetics:
 A/Gene: GDB:CST3
 A/Cross-references: GDB:119817; OMIM:105150
 A/Map position: 20p11.2-20p11.2
 A/Intons: 81/3; 119/3
 C/Superfamily: cystatin; cystatin homology
 C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline
 F/1-26/Domain: signal sequence #status predicted <SIG>
 F/27-146/Product: cystatin C #status experimental <MAT>
 F/35-146/Domain: cystatin homology <CYS>
 F/81-85/Region: inhibitory #status predicted
 F/82/Modified site: hydroxyproline (Pro) (partial) #status experimental
 F/93-109,123-143/Disulfide bonds: #status experimental

Query Match 37.0%; Score 94; DB 1; Length 146;
 Best Local Similarity 39.5%; Pred. No. 5e-05;
 Matches 17; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRVLKVGQRTDHLBYHLNVEWMTTCCKPE 43
 DB 61 NKASNDWYHSRALQVBARQIVGVNVPFLDVELGRRTCTCKTQ 103

RESULT 8
 Cystatin SA precursor - human
 C/Species: Homo sapiens (man)
 C/Date: 31-Mar-1989 #sequence revision 30-Jun-1989 #text_change 16-Jul-1999
 C/Accession: B29632; S02490; A41422; B27015
 R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.
 Gene 61, 329-338, 1987
 A/Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three members of the human cysteine-proteinase inhibitor family
 A/Reference number: A91589; MUID:88185836; PMID:3446578
 A/Accession: B29632
 A/Molecule type: DNA
 A/Residues: 1-141 <SAI>
 A/Cross-references: GB:M19673; GB:M19170; NID:g186403; PIDN:AAA3116.1; PID:g386826
 A/Note: the authors translated the codon GAC for residue 129 as Asn
 R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.
 Biol. Chem. Hoppe-Seyler 369, 191-197, 1988
 A/Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily distinct from family I
 A/Reference number: S02489; MUID:89076505; PMID:3202964
 A/Accession: S02490
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 21-141 <SA2>
 R/Isemura, S.; Saitoh, E.; Sanada, K.
 J. Biochem. 102, 693-704, 1987
 A/Title: Characterization and amino acid sequence of a new acidic cysteine proteinase in human saliva
 A/Reference number: A41422; MUID:88139220; PMID:3436950
 A/Accession: A41422
 A/Molecule type: protein
 A/Residues: 25-141 <ISE>
 R/Isemura, S.; Saitoh, E.; Sanada, K.; Isemura, M.; Ito, S.
 in Cysteine Proteinases and Their Inhibitors, Turk, V., ed., pp.497-505, Walter de Gruyter, Berlin, 1990
 A/Title: Cystatin S and the related cysteine proteinase inhibitors in human saliva.
 A/Reference number: A27015
 A/Accession: B27015
 A/Molecule type: protein
 A/Residues: 25-134, 'D', 136-141 <IS2>
 C/Genetics:
 A/Gene: GDB:CST2

A/Cross-references: GDB:119816; OMIM:123856
 A/Map position: 20p11.2-20p11.2
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 27-49/106-146 <BER>
 C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fluid patients with certain autoimmune diseases.
 C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an important role in the regulation of the immune response.
 C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorrhage with amyloidosis (Finnish type).
 C/Genetics:
 A/Gene: GDB:CST3
 A/Cross-references: GDB:119817; OMIM:105150
 A/Map position: 20p11.2-20p11.2
 A/Intons: 81/3; 119/3
 C/Superfamily: cystatin; cystatin homology
 C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline
 F/1-26/Domain: signal sequence #status predicted <SIG>
 F/27-146/Product: cystatin C #status experimental <MAT>
 F/35-146/Domain: cystatin homology <CYS>
 F/81-85/Region: inhibitory #status predicted
 F/82/Modified site: hydroxyproline (Pro) (partial) #status experimental
 F/93-109,123-143/Disulfide bonds: #status experimental

Query Match 33.1%; Score 84; DB 2; Length 141;
 Best Local Similarity 32.6%; Pred. No. 0.00099;
 Matches 14; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRVLKVGQRTDHLBYHLNVEWMTTCCKPE 43
 DB 56 NKATEDYRRLRLVRLARQIVGVNVPFLDVELGRRTCTCKSQ 98

RESULT 9
 A43428
 onchocystatin - nematode (Onchocerca volvulus)
 N/Alternate names: cysteine proteinase inhibitor; onchocerciasis antigen
 C/Species: Onchocerca volvulus
 C/Date: 04-Mar-1993 #sequence revision 18-Nov-1994 #text_change 17-Mar-2000
 C/Accession: A43428; B43927
 R/Lustigman, S.; Brozman, B.; Hulma, T.; Prince, A.M.; McKerrow, J.H.
 J. Biol. Chem. 267, 17339-17346, 1992
 A/Title: Molecular cloning and characterization of onchocystatin, a cysteine proteinase inhibitor from the nematode Onchocerca volvulus
 A/Reference number: A43428; MUID:92381053; PMID:1512269
 A/Accession: A43428
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-162 <LUS>
 A/Cross-references: GB:M37105; NID:g159905; PID:g159906
 A/Note: sequence extracted from NCBI backbone (NCBIN:111962, NCBI:P:111963)
 R/Chandrasekar, R.; Masood, K.; Alvarez, R.M.; Oguntirade, A.F.; Lujan, R.; Richards Jr
 J. Clin. Invest. 88, 1460-1466, 1991
 A/Title: Molecular cloning and characterization of recombinant parasite antigens for immunodiagnosis of onchocerciasis
 A/Reference number: A43927; MUID:92042729; PMID:1840605
 A/Accession: B43927
 A/Molecule type: mRNA
 A/Residues: 'P', 37-57, 'A', 59-71, 'R', 73-83, 'N', 85-126, 'W', 128-162 <CHA>
 A/Residues: 'P', 37-57, 'A', 59-71, 'R', 73-83, 'N', 85-126, 'W', 128-162 <CHA>
 A/Cross-references: GB:M60279; NID:g159888
 A/Experimental source: clone OC 9.3
 A/Note: sequence extracted from NCBI backbone (NCBIN:65111, NCBI:P:65113)
 C/Superfamily: cystatin; cystatin homology
 C/Keywords: cysteine proteinase inhibitor

Query Match 31.9%; Score 81; DB 2; Length 162;
 Best Local Similarity 34.1%; Pred. No. 0.0029;
 Matches 14; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRVLKVGQRTDHLBYHLNVEWMTTCCK 41
 DB 77 NEQSDNDEYHIMPILKLVSSQVAGVYKMDVQVARSQCKK 117

RESULT 10
 A28793
 cystatin - puff adder
 C/Species: Bitis arietans (puff adder)
 C/Date: 15-Dec-1988 #sequence revision 15-Dec-1988 #text_change 30-Sep-1993
 C/Accession: A28793
 R/Ritton, A.; Evans, H.J.; Machleidt, W.; Barrett, A.J.
 Biochem. J. 246, 799-802, 1987
 A/Title: Amino acid sequence of a cystatin from venom of the African puff adder (Bitis arietans)
 A/Reference number: A28793; MUID:88076861; PMID:3500714
 A/Accession: A28793
 A/Molecule type: protein
 A/Residues: 1-111 <RIT>
 C/Superfamily: cystatin; cystatin homology

Query Match 31.5%; Score 80; DB 2; Length 111;
 Best Local Similarity 39.0%; Pred. No. 0.0025;
 Matches 16; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKYHFRIFRVLKVGQRTDHLBYHLNVEWMTTCCK 41

Db 27 NAGSKNDYFRKRRVVAQSGVGVKXYLMELKTKCK 67

RESULT 11

JC4536
Cysteatin precursor - horseshoe crab (Trachyleus tridentatus)
N/Alternate names: cystatin L; cysteine proteinase inhibitor
C/Species: Trachyleus tridentatus
C/Date: 15-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 03-Dec-1999
C/Accession: JC4536; PC4122
R/Agarwala, K.L.; Kawabata, S.; Hirata, M.; Miyagi, M.; Tsumasawa, S.; Iwanaga, S.
J. Biochem. 119, 85-94, 1996
A/Title: A cysteine protease inhibitor scored in the large granules of horseshoe crab he
A/Reference number: JC4536; MUID:9706312; PMID:8907180
A/Accession: JC4536
A/Molecule type: mRNA
A/Residues: 1-133 <AGA>
A/Accession: PC4122
A/Molecule type: protein
A/Residues: 26-57/60-105/107-113/115-128 <AG2>
A/Experimental source: hemocytes
C/Comment: This protein belongs to the cystatin family and it is a single-chain protein
activity against Gram-negative bacteria, defense against invading microbes, and response to
C/Superfamily: cystatin; cystatin homology
C/Keywords: cysteine proteinase inhibitor; hemolymph; pyroglutamic acid
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-133/Product: limulus-cystatin #status predicted <MAT>
F/21-132/Domain: cystatin homology <CYS>
F/20/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F/85-98,109-129/Disulfide bonds: #status predicted

Query Match 29.5%; Score 75; DB 2; Length 133;
Best Local Similarity 30.2%; Pred. No. 0.014;
Matches 13; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Qy 1 NKESDDKYHFRIRPVLRKVRQVTDHLEHLNVEQMWTTCCKP 43
Db 47 SRSNSLYHKKLTKIKRATQVSGINVEYFIETGTTCKSE 89

RESULT 12

UDHUP2
Cystatin SN precursor (validated) - human
N/Alternate names: Cystatin SA-I
C/Species: Homo sapiens (man)
C/Date: 28-May-1986 #sequence_revision 08-Feb-1996 #text_change 08-Dec-2000
C/Accession: A28110; S02489; R29632; A01273; S19279
R/Al-Hashimi, I.; Dickinson, D.P.; Levine, M.J.
J. Biol. Chem. 263, 9381-9387, 1988
A/Title: Purification, molecular cloning, and sequencing of salivary cystatin, SA-I.
A/Reference number: A28110; MUID:88243825; PMID:2837486
A/Accession: A28110
A/Molecule type: mRNA
A/Residues: 1-141 <ALH>
A/Cross-references: GB:03870; NID:g337751; PID:AAA60299.1; PID:g337752
R/Saitoh, E.; Isemura, S.; Sanada, K.; Kim, H.S.; Smithies, O.; Maeda, N.
Biol. Chem. Hoppe-Seyler 369, 191-197, 1988
A/Title: Cystatin superfamily. Evidence that family II cystatin genes are evolutionarily
A/Reference number: S02489; MUID:89076505; PMID:3202964
A/Accession: S02489
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 21-141 <SA2>
R/Saitoh, E.; Kim, H.S.; Smithies, O.; Maeda, N.
Gene 61, 329-338, 1987
A/Title: Human cysteine-proteinase inhibitors: nucleotide sequence analysis of three mem
A/Reference number: A91589; MUID:88185835; PMID:3446578
A/Accession: A29632
A/Molecule type: DNA
A/Residues: 1-86, '1', 88-141 <SA1>
R/Isemura, S.; Saitoh, E.; Sanada, K.
FEBS Lett. 198, 145-149, 1986

A/Title: Characterization of a new cysteine proteinase inhibitor of human saliva, cystati
A/Reference number: A01273; MUID:86164938; PMID:3514272

A/Accession: A01273
A/Molecule type: protein
A/Residues: 29-141 <ISE>
R/Ramasubbu, N.; Reddy, M.S.; Bergey, E.J.; Harasathy, G.G.; Soni, S.D.; Levine, M.J.
Biochem. J. 280, 341-352, 1991
A/Title: Large-scale purification and characterization of the major phosphoproteins and
A/Reference number: S19279; MUID:92082465; PMID:11747107
A/Accession: S19279
A/Status: preliminary
A/Molecule type: protein
A/Residues: 21-55 <RAM>
C/Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a
ences. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl
C/Genetics:
A/Gene: GDB:CS71
A/Cross-references: GDB:119815; OMIM:123855
A/Map position: 20p11.2-20p11.2
C/Superfamily: cystatin; cystatin homology
C/Keywords: cysteine proteinase inhibitor; extracellular protein; saliva
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-141/Product: cystatin SA-I #status experimental <MAT>
F/29-141/Product: cystatin SN #status experimental <MAT2>
F/30-141/Domain: cystatin homology <CYS>
F/76-80/Region: inhibitory #status predicted
F/94-104,118-138/Disulfide bonds: #status predicted

Query Match 28.7%; Score 73; DB 1; Length 141;
Best Local Similarity 34.9%; Pred. No. 0.027;
Matches 15; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

Qy 1 NKESDDKYHFRIRPVLRKVRQVTDHLEHLNVEQMWTTCCKP 43
Db 56 NKATKDYRRPRLRARQQTVGVVFPDVEGRITCTKSQ 98

RESULT 13

JC2040
Cystatin - chum salmon
N/Alternate names: cysteine proteinase inhibitor
C/Species: Oncorhynchus keta (chum salmon)
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 16-Jul-1999
C/Accession: JC2040
R/Koide, Y.; Noso, T.
Biolci. Biotechnol. Biochem. 58, 164-169, 1994
A/Title: The complete amino acid sequence of pituitary cystatin from chum salmon.
A/Reference number: JC2040; MUID:94162738; PMID:7764512
A/Accession: JC2040
A/Molecule type: protein
A/Residues: 1-111 <KOI>
C/Comment: The intracellular role of this protein is the inhibition of intralysosomal pr
C/Superfamily: cystatin; cystatin homology
C/Keywords: cysteine proteinase inhibitor
F/2-111/Domain: cystatin homology <CYS>
F/48-52/Region: inhibitory
F/89-109/Disulfide bonds: #status experimental

Query Match 28.0%; Score 71; DB 1; Length 111;
Best Local Similarity 31.7%; Pred. No. 0.039;
Matches 13; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKESDDKYHFRIRPVLRKVRQVTDHLEHLNVEQMWTTCCK 41
Db 28 NKXTNDFVRQVAKVNAQKQVSGMKYFTVQMGTRPCRK 68

RESULT 14

JC4918
Cystatin precursor - chum salmon
C/Species: Oncorhynchus keta (chum salmon)
C/Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jan-2000
C/Accession: JC4918

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 5.47619 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-10

Sequence: 1 NKESDDKYHFRIVLKVQR.....EYHLNVEMQWTCQKPEPTN 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	254	100.0	137	CS11_HUMAN	Q9H12 homo sapien
2	162	63.8	139	CS11_MOUSE	Q9D69 mus musculu
3	118	46.5	142	CST8_HUMAN	O60676 homo sapien
4	111	43.7	142	CST8_MOUSE	P32766 mus musculu
5	106	41.7	142	CST8_RAT	O88968 ratu musculu
6	103	40.6	127	CYT_C_RAT	P14841 ratu musculu
7	101	39.8	148	CYT_C_RABIT	O97862 ratu musculu
8	99	39.0	140	CYT_C_MOUSE	P21460 mus musculu
9	98	38.6	148	CYT_C_BOVIN	P01035 bos taurus
10	96	37.8	146	CYT_C_SALIC	O19093 salmuri bci
11	94	37.0	139	CYT_C_CHICK	P01038 gallu galli
12	94	37.0	146	CYT_C_HUMAN	P01034 homo sapien
13	94	37.0	146	CYT_C_MACMU	O19092 macaca mula
14	90	35.4	116	CYT_COTTA	P01064 cotu musculu
15	87	34.3	165	CSTL_HUMAN	Q9H14 homo sapien
16	84	33.1	141	CYT_H_HUMAN	P09228 homo sapien
17	81	31.9	162	CYT_X_ONCVO	P22085 onchocece
18	80	31.5	111	CYT_B_BITAR	P08935 bitis ariet
19	78	30.7	149	CYT_M_HUMAN	P15828 homo sapien
20	76	29.9	129	CYT_C_YPCA	P35481 cyprinu ca
21	73	28.7	141	CYT_N_HUMAN	P01036 mus sapien
22	71	28.0	130	CYT_ONCKE	Q98967 oncothynu
23	71	28.0	130	CYT_ONCMY	Q91135 oncothynu
24	71	28.0	141	CYT_S_HUMAN	P01036 homo sapien
25	67	26.4	142	CYT_H_HUMAN	P28335 homo sapien
26	65	25.6	145	CYT_F_HUMAN	O76096 homo sapien
27	63	24.8	438	EXG_CANAL	P29717 candida alb
28	62	24.4	434	KNL2_BOVIN	P01047 bos taurus
29	62	24.4	619	KNH2_BOVIN	P01045 bos taurus
30	61	24.0	436	KNH1_BOVIN	P01046 bos taurus
31	61	24.0	621	KNH1_BOVIN	P01046 bos taurus
32	60	23.6	257	CG1C_ORYSA	P03411 cryza sattiv
33	59	23.2	644	KNH1_HUMAN	P01042 homo sapien

34	58.5	23.0	137	1	CST9_MOUSE	Q9Z0H6 mus musculu
35	58	22.8	144	1	CYT_F_MOUSE	O89098 mus musculu
36	58	22.8	448	1	EXG1_YEAST	P23776 saccharomyc
37	57.5	22.6	2200	1	LAR_CAEEL	Q9B0M8 caenorhabdi
38	57	22.4	141	1	CYT_S_RAT	P19313 ratu musculu
39	56	22.0	200	1	SP24_BOVIN	Q27967 bos taurus
40	56	22.0	621	1	MAI1_MAIZE	P17571 zea mays (m
41	56	22.0	661	1	KNH_MOUSE	O08677 mus musculu
42	55	21.7	527	1	YBIF_ECOLI	P75745 escherichia
43	55	21.7	862	1	Z409_HUMAN	Q9UP06 homo sapien
44	55	21.7	904	1	MAI1_TOBAC	P11605 nicotiana t
45	55	21.7	904	1	MAI2_TOBAC	P08509 nicotiana t

ALIGNMENTS

RESULT 1
ID CS11_HUMAN STANDARD; PRT; 137 AA.
AC Q9H12; Q9H13;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin 11 precursor.
GN CS11 OR CST8L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavridis G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Griffiths D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levasseur M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurtry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sultoni J.E.,
RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
-1- SUBCELLULAR LOCATION: Secreted (Potential).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q9H12-1; Sequence=Displayed;
Name=2;
IsoId=Q9H12-2; Sequence=VSP_001260;
Note=No experimental confirmation available;
-1- SIMILARITY: Belongs to the cystatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Ross M.T., Scott C.B., Senta H.K., Showstken R., Sims S.,
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin R., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF059244; AAC14707.1; -;
 CC EMBL: AL109954; CAB64234.1; -;
 CC HSSP: P01034; 1G96
 CC Genew: HGNC:2480; CST8
 CC GO: GO:0004869; F:cysteine protease inhibitor activity; TAS.
 CC InterPro: IPR000010; Cystatin.
 CC Pfam: PF00031; Cystatin; 1.
 CC SMART: SM00043; Cy; 1.
 CC Thiol protease inhibitor; Signal; Polymorphism.
 KW SIGNAL
 FT CHAIN 1 21 POTENTIAL.
 FT CHAIN 22 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 95 105 BY SIMILARITY.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 142 142 A -> P (in dbSNP:1054633).
 FT SIGNAL 1 /FTID=VAR 014527.
 SQ SEQUENCE 142 AA; 16275 MW; 9A512757E0F4E0C CRC64;
 Query Match 46.5%; Score 118; DB 1; Length 142;
 Best Local Similarity 50.0%; Pred. No. 1.4e-08;
 Matches 23; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
 Oy 1 NKESDDKYHFRIFRVLVKVRQVTDHLEHVLAVEMQWTTCKP 46
 Db 57 NKESDDKYHFRIFRVLVKVRQVTDHLEHVLAVEMQWTTCKP 102
 ID NKESDDKYHFRIFRVLVKVRQVTDHLEHVLAVEMQWTTCKP 46
 AC P32766; O89102; PRT; 142 AA.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-
 DE related epididymal specific protein) (Cystatin 8).
 GN CST8 OR CR8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H, and CD-1;
 RX MEDLINE=9247899; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.;
 RT "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermatogenic gene.";
 RL Biochem. J. 340:85-93(1999).
 RN [2]
 RP SEQUENCE OF 4-142 FROM N.A.
 RC TISSUE=Epididymis;
 RX MEDLINE=9307899; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;
 RT "The CR8 gene: a unique testis-regulated gene related to the cystatin
 RT family is highly restricted in its expression to the proximal region
 RT of the mouse epididymis.";
 RL Mol. Endocrinol. 6:1653-1664(1992).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- INDUCTION: Testicular factors or hormones other than androgens
 CC present in the testicular fluid may be involved in the regulation
 CC of CR8 gene expression.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF091503; AAC61754.1; -;
 CC EMBL: AF090691; AAC36316.1; -;
 CC EMBL: S49926; AAC35390.1; -;
 CC PIR: A45361; A45361.
 CC HSSP: P01034; 1G96.
 CC MGD: MGI:107161; Cst8.
 CC InterPro: IPR000010; Cystatin.
 CC Pfam: PF00031; Cystatin; 1.
 CC SMART: SM00043; Cy; 1.
 CC Thiol protease inhibitor; Signal.
 KW SIGNAL
 FT CHAIN 1 19 POTENTIAL.
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 95 105 BY SIMILARITY.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 4 15 PLSLTIPLIP -> GTRBOVGSQK (IN REF. 2).
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B39F6673E CRC64;
 Query Match 43.7%; Score 111; DB 1; Length 142;
 Best Local Similarity 42.9%; Pred. No. 1.2e-07;
 Matches 16; Conservative 13; Mismatches 11; Indels 0; Gaps 0;
 Oy 1 NKESDDKYHFRIFRVLVKVRQVTDHLEHVLAVEMQWTTCKP 42
 Db 57 NKESDDKYHFRIFRVLVKVRQVTDHLEHVLAVEMQWTTCKP 98
 ID NKESDDKYHFRIFRVLVKVRQVTDHLEHVLAVEMQWTTCKP 42
 AC O88969; PRT; 142 AA.
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
 DE 8).
 GN CST8 OR CRES.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;
 RX MEDLINE=9924789; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.;
 RT "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermatogenic gene";
 RL Biochem. J. 340:85-93(1999).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF090692; AAC36317.1; -.
 DR HSP: P01034; 1696.
 DR InterPro: IPR00010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; Cy; 1.
 KM Thiol protease inhibitor; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT PROTEIN.
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 95 105 BY SIMILARITY.
 FT DISULFID 119 139 BY SIMILARITY.
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 142 AA; 16246 MW; FB873FMA6BCAB34 CRC64;
 Query Match 41.7%; Score 106; DB 1; Length 142;
 Best Local Similarity 42.9%; Pred. No. 5.3e-07; Indels 0; Gaps 0;
 Matches 18; Conservative 11; Mismatches 13;
 QY 1 NKESDDKYHFRIPVLKVRQVTDHLEYNLWEMQWTTCKRP 42
 DB 57 NKGSDDKYLFLDKTLHATLQITDRMEHYIDVQISRSNCRKP 98
 RESULT 6
 CYTC RAT STANDARD; PRT; 127 AA.
 AC PI4841;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor (Fragment).
 GN CST3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Buffalo;
 RX MEDLINE=90092122; PubMed=2689174;
 RA Cole T., Dickson P.W., Ebnard F., Averill F., Risbrigger G.,
 RA Gauthier F., Schreiber G.;
 RT "The CDNA structure and expression analysis of the genes for the

RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
 RT in rat brain.";
 RL Eur. J. Biochem. 186:35-42(1989).
 RN [2]
 RP SEQUENCE OF 8-127.
 RX MEDLINE=90380276; PubMed=2400577;
 RA Ebnard F., Ebnard A., Faucher D., Capony J.-P., Derancourt J.,
 RA Brillard M., Gauthier F.;
 RT "Rat cystatin C: the complete amino acid sequence reveals a site for
 RT N-glycosylation";
 RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
 RN [3]
 RP SEQUENCE OF 8-49.
 RX MEDLINE=88313020; PubMed=3044831;
 RA Ebnard A., Ebnard F., Faucher D., Gauthier F.;
 RT "Two rat homologues of human cystatin C";
 RL FEBS Lett. 236:475-478(1988).
 RN [4]
 RP SEQUENCE OF 8-20.
 RC TISSUE=Sertoli cells;
 RX MEDLINE=9225121; PubMed=1563513;
 RA Ebnard A., Ebnard F., Guillo F., Gauthier F.;
 RT "Production of the cysteine proteinase inhibitor cystatin C by rat
 RT Sertoli cells";
 RL FEBS Lett. 300:131-135(1992).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity. Known to inhibit cathepsin B,
 CC H. and L.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16957; CA334831.1; -.
 DR PIR: S07085; S07085.
 DR PIR: S10587; S10587.
 DR HSP: P01034; 1696.
 DR InterPro: IPR00010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; Cy; 1.
 KM Thiol protease inhibitor; Signal.
 FT NON TER 1 1
 FT CHAIN 8 127 CYSTATIN C.
 FT ACT_SITE 18 18 REACTIVE SITE.
 FT SITE 62 66 SECONDARY AREA OF CONTACT.
 FT DISULFID 80 90 BY SIMILARITY.
 FT DISULFID 104 124 BY SIMILARITY.
 FT CONFLICT 25 25 A -> E (IN REF. 2).
 SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;
 Query Match 40.6%; Score 103; DB 1; Length 127;
 Best Local Similarity 43.2%; Pred. No. 1.2e-06;
 Matches 19; Conservative 11; Mismatches 14; Indels 0; Gaps 0;
 QY 1 NKESDDKYHFRIPVLKVRQVTDHLEYNLWEMQWTTCKRP 44
 DB 42 NKGSNDAYHSPRAIQVVARQOLVAGINYLVDWEGRTTCTKSOT 85
 RESULT 7
 CYTC RABIT STANDARD; PRT; 148 AA.
 AC 097862;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor.
 GN CST3.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN NCBI_TaxID=9986;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Japanese white; TISSUE=Bone;
 RX MEDLINE=98424349; PubMed=9753427;
 RA Kobori M., Ikeda Y., Nara H., Kamegawa M., Nojima H.,
 RA Kawahara H.;
 RT "Large scale isolation of osteoclast-specific genes by an improved
 RT method involving the preparation of a subtracted cDNA library."
 RL Gene 3:459-475(1998).
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB009342; BAA75921.1; -
 CC HSSP; P01034; 1G96.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART; SM00043; CY 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 DR Thiol protease inhibitor; Signal.
 KW SIGNAL 1 28
 FT CHAIN 29 148 POTENTIAL.
 FT ACT SITE 39 39 CYSTATIN C.
 FT SITE 83 87 REACTIVE SITE.
 FT DISULFID 101 111 SECONDARY AREA OF CONTACT.
 FT DISULFID 125 145 BY SIMILARITY.
 SQ SEQUENCE 148 AA; 16346 MW; 1523C8311695B9A CRC64;
 Query Match 39.8%; Score 101; DB 1; Length 148;
 Best Local Similarity 40.9%; Pred. No. 2.5e-05;
 Matches 18; Conservative 13; Mismatches 13; Indels 0; Gaps 0;
 QY 1 NKESDDKXHPRIPIVLRVQKQVTDHLEYNVEMQTTCKPPT 44
 DB 63 NKGSNDRYHRAIQVVARARQIVSGVYVYDLIGRTTCKTQTP 106
 RESULT 8
 CYTC MOUSE STANDARD; PRT; 140 AA.
 ID CYTC MOUSE STANDARD; PRT; 140 AA.
 AC P21460;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin C precursor (Cystatin 3).
 GN CST3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE=91054522; PubMed=2241883;
 RA Solem M., Rawson C., Lindburg K., Barnes D.;
 RT "Transforming growth factor beta regulates cystatin C in serum-free
 RT mouse embryo (SPME) cells."
 RL Biochem. Biophys. Res. Commun. 172:945-951(1990).
 [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv; TISSUE=Liver;
 RX MEDLINE=95137392; PubMed=7835704;
 RA Hun C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
 RT "Structural organization, expression and chromosomal mapping of the
 RT mouse cystatin-C-encoding gene (Csc3)."
 RL Gene 152:221-226(1995).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ILS, and ISB;
 RX MEDLINE=21363810; PubMed=11471062;
 RA Eisinger M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
 RT "High-throughput sequence identification of gene coding variants
 RT within alcohol-related OTRs."
 RL Mamm. Genome 12:657-663(2001).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uden T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59470; AAA63298.1; -
 CC EMBL; U10098; AAB41056.1; -
 CC EMBL; AF483486; AAL90760.1; -
 CC EMBL; AF483487; AAL90761.1; -
 CC EMBL; BC002072; AAH02072.1; -
 CC PIR; A36163; A36163.
 DR HSSP; P01034; 1G96.
 DR MGD; MGI:102519; Cat3.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR SMART; SM00043; CY 1.
 KW Thiol protease inhibitor; Signal.
 KW SIGNAL 1 20
 FT CHAIN 21 140
 FT ACT SITE 31 31 CYSTATIN C.
 FT SITE 75 79 REACTIVE SITE.
 FT DISULFID 93 103 SECONDARY AREA OF CONTACT.
 FT DISULFID 117 137 BY SIMILARITY.
 FT CONFLICT 16 16 A -> G (IN REF. 1).
 FT CONFLICT 84 84 L -> F (IN REF. 1).

SQ SEQUENCE 140 AA; 15531 MW; 3A563406DD58D0F5 CRC64;

Query Match 39.0%; Score 99; DB 1; Length 140;
Best Local Similarity 43.2%; Pred. No. 4.4e-06;
Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDKYHFRIFRVLKVQRQVTDHLEHYNVEMQWTTCKRPE 44
Db 55 NKGSNDAYHSRAIOVVARAKQIVAGVNYFLDVEEMGRITTCRKSQT 98

RESULT 9

CYT_C_BOVIN STANDARD; PRT; 148 AA.

AC P01035;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).
GN CST3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.
RC TISSUE=Cerebrospinal fluid, and Choroid plexus;
RX MEDLINE=98094199; PubMed=9434110;
RA Olsson S.-L., Ek B., Wilm M., Broberg S., Rask L., Bjorker I.;
RT "Molecular cloning and N-terminal analysis of bovine cystatin C
identification of a full-length N-terminal region.";
RL Biochim. Biophys. Acta 1343:203-210(1997).
[2]
RP SEQUENCE OF 37-148.
RX MEDLINE=85231205; PubMed=3891407;
RA Hirado M., Tsunashima S., Sakiyama F., Ninobe M., Fujii S.;
RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine
proteinase inhibitor.";
RL FEBS Lett. 186:41-45(1985).
CC -1- FUNCTION: This is a thiol proteinase inhibitor.
CC -1- MASS SPECTROMETRY: MW=13420, METHOD=MALDI.
CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: Y10811; CAA71771.1; -.
DR HSSP: P01034; I966.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin.1.
DR SMART: SM00043; CY.1.
DR PROSITE: PS00287; CYSTATIN.1.
KW Thiol protease inhibitor; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 30 PROBABLE.
FT CHAIN 31 148 CYSTATIN C.
FT MOD_RES 31 31 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT ACT_SITE 40 40 REACTIVE SITE.
FT SITE 84 88 SECONDARY AREA OF CONTACT.
FT DISULFID 102 112 BY SIMILARITY.
FT DISULFID 126 146 BY SIMILARITY.
SQ SEQUENCE 148 AA; 16265 MW; EE740FE37CFB9F0E CRC64;

Query Match 38.6%; Score 98; DB 1; Length 148;
Best Local Similarity 41.9%; Pred. No. 6.3e-06;
Matches 18; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDKYHFRIFRVLKVQRQVTDHLEHYNVEMQWTTCKRPE 43

Db 64 NKGSNDAYHSRAIOVVARAKQIVAGVNYFLDVEEMGRITTCRKSQT 106

RESULT 10

CYT_C_SALSC STANDARD; PRT; 146 AA.

AC O19093;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Cystatin C precursor.
GN CST3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
NCBI_TaxID=9521;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
stroke 27:2080-2085(1996)."
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U52028; AAB64051.1; -.
DR HSSP: P01034; I966.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin.1.
DR SMART: SM00043; CY.1.
DR PROSITE: PS00287; CYSTATIN.1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;

Query Match 37.8%; Score 96; DB 1; Length 146;
Best Local Similarity 43.9%; Pred. No. 1.1e-05;
Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDKYHFRIFRVLKVQRQVTDHLEHYNVEMQWTTCKR 41
Db 61 NKASNDYHSRAIOVVARAKQIVAGVNYFLDVEEMGRITTCR 101

RESULT 11

CYT_CHICK STANDARD; PRT; 139 AA.

AC P01038;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Cystatin precursor (Egg-white cystatin).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeopteryx; Aves; Neognathae; Galliformes; Phasianidae; Gallinae.

OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A. MEDLINE=9008873; PubMed=2793849;
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
 RT and tissue distribution.";
 RL J. Biol. Chem. 264:17164-17169(1989).
 RN [2]
 RP SEQUENCE OF 24-139. MEDLINE=84178305; PubMed=6712597;
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
 RT "Cystatin. Amino acid sequence and possible secondary structure.";
 RL Biochem. J. 217:813-817(1984).
 RN [3]
 RP SEQUENCE OF 24-139. MEDLINE=84110059; PubMed=6662498;
 RA Turk V., Brzin J., Longer M., Ritonja A., Brotkin M., Borchart U.,
 RA Machleidt W.;
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 RT of cystatin from chicken egg white.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
 RN [4]
 RP CHARACTERIZATION OF PROTEIN. MEDLINE=83256421; PubMed=6409085;
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
 RA Senter D.C., Barrett A.J.;
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
 RT purification from egg white, characterization, and detection in
 RT chicken serum.";
 RL Biochem. J. 211:129-138(1983).
 RN [5]
 RP DISULFIDE BONDS. Grubb A., Loebberg H., Barrett A.J.;
 RA "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 RT cystatin.";
 RL FEBS Lett. 170:370-374(1984).
 RN [6]
 RP PHOSPHORYLATION. MEDLINE=89252033; PubMed=2721673;
 RA Leber W., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
 RA Bode W.;
 RT "The cysteine proteinase inhibitor chicken cystatin is a
 RT phosphoprotein.";
 RL FEBS Lett. 248:162-168(1989).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS). MEDLINE=89052676; PubMed=3191914;
 RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,
 RA Brzin J., Kos J., Turk V.;
 RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
 RT its possible mode of interaction with cysteine proteinases.";
 RL EMBO J. 7:2593-2599(1988).
 RN [8]
 RP STRUCTURE BY NMR. MEDLINE=94087719; PubMed=8263912;
 RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,
 RA Auerwald E.A., Jeanlike R., Oeschkinat H.;
 RT "The structures of native phosphorylated chicken cystatin and of a
 RT recombinant unphosphorylated variant in solution.";
 RL J. Mol. Biol. 234:1048-1059(1993).
 CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of
 CC thiol proteases including ficin, papain, and cathepsins B, C, H,
 CC and L. Although isolated from egg white, it is also present in
 CC serum.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; J05077; AAA48744.1; -.
 DR PIR; A34456; UDCH.
 DR PDB; 1CEW; 31-JAN-94.
 DR PDB; 1A67; 27-MAY-98.
 DR PDB; 1A90; 17-JUN-98.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 139
 FT ACT_SITE 32 32
 FT SITE 76 80
 FT DISULFID 94 104
 FT DISULFID 118 138
 FT MOD_RES 103 103
 FT STRAND 35 36
 FT TURN 39 40
 FT TURN 42 51
 FT HELIX 42 51
 FT TURN 52 52
 FT HELIX 53 56
 FT TURN 57 58
 FT STRAND 63 77
 FT STRAND 81 95
 FT TURN 96 97
 FT TURN 99 100
 FT HELIX 101 108
 FT STRAND 115 125
 FT TURN 126 129
 FT STRAND 130 139
 SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;
 Query Match 37.0%; Score 94; DB 1; Length 139;
 Best Local Similarity 43.9%; Pred. No. 2e-05;
 Matches 18; Conservative 9; Mismatches 14; Indels 0; Gaps 0;
 Oy 1 NKESDDKHYHFRIFRYLVKQROVTDHLYHNLVNEKQMTTCOK 41
 Db 56 NRASNDKYSRRVVRVIAKROLVSGIKYILQVEIGRTTCPK 96
 RESULT 12
 ID CYTC HUMAN STANDARD; PRT; 146 AA.
 AC P01034;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
 DE (Post-gamma-globulin).
 GN CST3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. MEDLINE=87219149; PubMed=3495457;
 RC TISSUE=Placenta;
 RX MEDLINE=87219149; PubMed=3495457;
 RA Abrahamsen M., Grubb A., Olafsson I., Lundwall A.;
 RT "Molecular cloning and sequence analysis of cDNA coding for the
 RT precursor of the human cysteine proteinase inhibitor cystatin C.";
 RL FEBS Lett. 216:229-233(1987).
 RN [2]
 RP SEQUENCE FROM N.A. MEDLINE=90303202; PubMed=2363674;
 RC TISSUE=Leukocyte;
 RX MEDLINE=90303202; PubMed=2363674;
 RA Abrahamsen M., Olafsson I., Palsdottir A., Ulfvaseck M., Lundwall A.,
 RA Jensen O., Grubb A.;
 RT "Structure and expression of the human cystatin C gene.";

RL Biochem. J. 268:287-294(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (HCHWA VARIANT).
 RC TISSUE=Brain;
 RX MEDLINE=89235594; PubMed=2541223;
 RA Levy E., Lopez-Olin C., Ghiso J., Gellner D., Frangione B.;
 RT "stroke in Icelandic patients with hereditary amyloid angiopathy is
 RT related to a mutation in the cystatin C gene, an inhibitor of
 RT cysteine proteases.";
 RL J. Exp. Med. 169:1771-1778(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89350949; PubMed=2764935;
 RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,
 RA Iemura S., Sanada K.;
 RT "The human cystatin C gene (CST3) is a member of the cystatin gene
 RT family which is localized on chromosome 20.";
 RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;
 RT "Acquisition of complex patterns of differential expression in
 RT epithelial cell populations during the evolution of type 2 cystatin
 RT genes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Barker A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor P.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharvestho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.I., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.V., Ramsay H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showstken R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Uebei T.B., Toshiyuki S., Carrinci F., Prange C.,
 RA Raha S.S., Loquailano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Kizylinski M.I., Skalek A., Smalins D.E.,
 RA Schercher A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 27-146.
 RX MEDLINE=82222268; PubMed=6283552;
 RA Grubb A., Loeffberg H.;
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and
 RT presence in the adenohypophysis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).
 RN [9]
 RP SEQUENCE OF 27-73.
 RX MEDLINE=84110059; PubMed=6662498;
 RA Turk V., Brzin J., Longner M., Ritonja A., Eropkin M., Borchart U.,
 RA Machleidt W.;
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 RT of cystatin from chicken egg white.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
 RN [10]
 RP SEQUENCE OF 27-76.
 RX MEDLINE=84128015; PubMed=6365094;
 RA Brzin J., Popovic T., Turk V.;
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases.";
 RL Biochem. Biophys. Res. Commun. 118:103-109(1984).
 RN [11]
 RP DISULFIDE BONDS.
 RA Grubb A., Loeffberg H., Barrett A.J.;
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 RT cystatin.";
 RL FEBS Lett. 170:370-374(1984).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.
 RX MEDLINE=91173909; PubMed=11276250;
 RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,
 RA Abrahamson M., Jaskolski M.;
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through
 RT three-dimensional domain swapping.";
 RL Nat. Struct. Biol. 8:316-320(2001).
 RN [13]
 RP VARIANT Gln-94.
 RX MEDLINE=92316504; PubMed=1352269;
 RA Abrahamson M., Jonasson S., Olafsson I., Jensen O., Grubb A.;
 RT "Hereditary cystatin C amyloid angiopathy: identification of the
 RT disease-causing mutation and specific diagnosis by polymerase chain
 RT reaction-based analysis.";
 RL Hum. Genet. 89:377-380(1992).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SUBUNIT: Homodimer.
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,
 CC vas deferens, brain, thymus, and ovary and the lowest in the
 CC submandibular gland.
 CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral
 CC hemorrhage with amyloidosis (HCHWA) [MIM:105150]; also known as
 CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis
 CC Icelandic type. HCHWA is characterized by a thickening of the
 CC cerebral arteries walls with deposition of material with the
 CC characteristics of amyloid.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL, X05607; CAA29096.1; -


```
DR EMBL: X52255; CAA36497.1; -.
DR EMBL: M27891; AAA52164.1; -.
DR EMBL: M27889; AAA52164.1; JOINED.
DR EMBL: M27890; AAA52164.1; JOINED.
DR EMBL: X61681; CAA43856.2; -.
DR EMBL: X61682; CAA43856.2; JOINED.
DR EMBL: X61683; CAA43856.2; JOINED.
DR EMBL: A919564; AAK1570.1; -.
DR EMBL: AL121894; CAC05424.1; -.
DR EMBL: BC013083; AAI13083.1; -.
DR PIR: S10216; UDHU.
DR PDB: 1G96; 06-APR-01.
DR Genew: HGNC:2475; CST3.
DR MIM: 604312; -.
DR MIM: 105150; -.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
DR Thiol protease inhibitor; Amyloid; Signal; Disease mutation.
KW Polymorphism; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 146
FT ACT SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109
FT DISULFID 123 143

Query Match 37.0%; Score 94; DB 1; Length 146;
Best Local Similarity 39.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKESDDKYHPRIFRVLKVRQVTDHLEHNVEMQWTTCKPK 43
Db 61 NKASNDMYHSRALQVBARQIVAGVYFIDVLEGRITTKTKQ 103

RESULT 13
CYTC_MACNU STANDARD; PRT; 146 AA.
AC 015092;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
DE Cystatin C precursor.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; Pubmed=8998820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C, Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
-1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
thought to serve an important physiological role as a local
regulator of this enzyme activity.
-1- SIMILARITY: Belongs to the cystatin family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
-----
CC EMBL: U51912; AAB64050.1; -.
CC HSSP: P01034; 1G96.
```

```
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26
FT CHAIN 27 146
FT ACT SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109
FT DISULFID 123 143
SQ SEQUENCE 146 AA; 15857 MW; F0B3BB774A29DF26 CRC64;

Query Match 37.0%; Score 94; DB 1; Length 146;
Best Local Similarity 39.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKESDDKYHPRIFRVLKVRQVTDHLEHNVEMQWTTCKPK 43
Db 61 NKASNDMYHSRALQVBARQIVAGVYFIDVLEGRITTKTKQ 103

RESULT 14
CYT_COTUA STANDARD; PRT; 116 AA.
AC P81061;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin (Bgg-white cystatin).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE.
RC TISSUE=Bgg white;
RX MEDLINE=97420480; Pubmed=9276465;
RA Gerhartz B., Engh R.A., Mentele R., Eckerskorn C., Torguato R.,
RA Witman J., Kolb H.J., Machleidt W., Fritz H., Auerwald E.A.;
RT "Quail cystatin: isolation and characterization of a new member of
RT the cystatin family and its hypothetical interaction with cathepsin
RT B.";
RL FBS Lett. 412:551-558(1997).
-1- FUNCTION: This protein binds tightly to and inhibits papain and
cathepsin B.
-1- SIMILARITY: Belongs to the cystatin family.
CC HSSP: P01038; ICEW.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Phosphorylation.
FT ACT SITE 9 9 REACTIVE SITE.
FT SITE 53 57
FT DISULFID 71 81
FT DISULFID 95 115
FT MOD RES 80 80
SQ SEQUENCE 116 AA; 13093 MW; 4824862105A2AF70 CRC64;

Query Match 35.4%; Score 90; DB 1; Length 116;
Best Local Similarity 39.0%; Pred. No. 5.5e-05;
Matches 16; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 1 NKESDDKYHPRIFRVLKVRQVTDHLEHNVEMQWTTCKPK 41
Db 33 NKASNDKYSSRVYRISAKQIVSGIKYIMEVIGRTTCKP 73

RESULT 15
CSTL_HUMAN STANDARD; PRT; 165 AA.
```

AC 09H114;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-like 1 precursor.
 GN CSTL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehesaialho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie I.J., McElay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showlken R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromens A.C., Vautin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:665-871(2001).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: AL096677; CAC03440.2; -
 DR HSSP; P01038; 1A90.
 DR Genew; HGNC:15958; CSTL1.
 DR Interpro; IPR000010; Cystatin.
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.
 KM Thiol protease inhibitor; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 165 CYSTATIN-LIKE 1.
 FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 111 121 BY SIMILARITY.
 FT DISULFID 134 154 BY SIMILARITY.
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 165 AA; 19312 MW; 9D6D685875DAEEA CRC64;

Query Match 34.3%; Score 87; DB 1; Length 165;
 Best Local Similarity 27.3%; Pred. No. 0.0002;

Matches 18; Conservative 15; Mismatches 13; Indels 20; Gaps 1;
 QY 1 NKESDDKYPRIIRVILKQ-----RQVTDHLEHLNVEMQTTCQ 40
 DB 53 NNASNDIYLRYVQRLIRSQWQERVSHMLGVHINSTDSQRLTGVGVYIVTKIGWKCK 112
 QY 41 KPEITN 46
 DB 113 RNDTSN 118
 Search completed: March 18, 2004, 14:16:11
 Job time : 5.47619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 27.5635 Seconds
(without alignments)
526.560 Million cell updates/sec

Title: US-09-941-314-10

Perfect score: 254
Sequence: 1 NKSDDKXHFRIKRVLR.....EYHNVEMQWTTCKPPTN 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	138	4	Q8WXU6
2	151	59.4	139	11	Q8K5A3
3	113	44.5	141	11	Q9DAP1
4	113	44.5	141	11	Q80ZM5
5	110	43.3	103	4	Q8WXU5
6	99	39.0	140	11	Q9BFX9
7	92	36.2	112	13	Q9BSR4
8	92	36.2	112	13	Q9BSR3
9	81	31.9	109	5	Q9TV65
10	81	31.9	148	5	Q9NH95
11	80	31.5	161	5	Q16159
12	77	30.3	140	11	Q80Y72
13	76	29.9	425	3	Q12700
14	72	28.3	125	5	Q25620
15	72	28.3	146	11	Q8K397
16	72	28.3	149	11	Q9D1B1

17	71	28.0	149	11	Q8VHC1	Q8VHC1 ratnus norv
18	71	28.0	498	5	Q16454	Q16454 caenorhabdi
19	68	26.8	157	5	Q17108	Q17108 acanthochel
20	67	26.4	148	11	Q9JH84	Q9JH84 mus musculu
21	65	25.6	167	4	Q724U8	Q724U8 homo sapien
22	64	25.2	144	13	Q8JFUS	Q8JFUS brachydanio
23	64	25.2	284	16	Q88Z07	Q88Z07 lactobacilli
24	63	24.8	438	3	Q9URL8	Q9URL8 candida alb
25	62	24.4	302	4	Q9H740	Q9H740 homo sapien
26	62	24.4	421	4	Q9NKS0	Q9NKS0 homo sapien
27	62	24.4	693	11	Q8K145	Q8K145 mus musculu
28	62	24.4	995	4	Q9HC18	Q9HC18 homo sapien
29	61.5	24.2	587	9	Q85911	Q85911 staphylococ
30	61.5	24.2	587	9	Q85910	Q85910 staphylococ
31	61	24.0	127	5	P90598	P90598 brugia mala
32	61	24.0	787	11	Q8B182	Q8B182 mus musculu
33	60	23.6	204	4	Q8TCY7	Q8TCY7 homo sapien
34	60	23.6	275	4	Q8WYG2	Q8WYG2 homo sapien
35	60	23.6	357	4	Q72465	Q72465 homo sapien
36	59	23.2	429	3	Q12628	Q12628 kiuyveromyc
37	58.5	23.0	133	5	Q8WVB6	Q8WVB6 ixodes scap
38	58.5	23.0	246	16	Q980Y6	Q980Y6 mycoplasma
39	58.5	23.0	544	6	Q28548	Q28548 ovis aries
40	58.5	23.0	891	5	Q81B61	Q81B61 plasmodium
41	58	22.8	159	4	Q8TDS3	Q8TDS3 homo sapien
42	58	22.8	167	11	Q9OWL5	Q9OWL5 mus musculu
43	58	22.8	540	5	Q8T9G8	Q8T9G8 drosophila
44	58	22.8	610	4	Q9UJ02	Q9UJ02 homo sapien
45	58	22.8	631	4	Q9NXK3	Q9NXK3 homo sapien

ALIGNMENTS

RESULT 1	Q8WXU6	PRELIMINARY;	PRT;	138 AA.
ID	Q8WXU6			
AC	Q8WXU6			
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)			
DE	SC13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.;			
RT	"SC13: A novel epididymal specific member of the cystatin family."			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF35480; AAL7191.1; -			
DR	GO; GO:0004869; P:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; CY; 1.			
DR	SEQUENCE	138 AA;	16506 MW;	E49440ACA3585C64 CRC64;
SO	SEQUENCE			
Query Match	100.0%;	Score 254;	DB 4;	Length 138;
Best local Similarity	100.0%;	Pred. No. 1.8e-26;		
Matches	46;	Conservative 0;	Mismatches 0;	Gaps 0;
OY	1 NKSDDKXHFRIKRVLRQVTDHLEHYHNVEMQWTTCKPPTN 46			
Db	56 NKSDDKXHFRIKRVLRQVTDHLEHYHNVEMQWTTCKPPTN 101			
RESULT 2	Q8K5A3	PRELIMINARY;	PRT;	139 AA.
ID	Q8K5A3			
AC	Q8K5A3			
DT	01-OCT-2002 (TREMBlrel. 22, Created)			
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)			

DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE Cystatin 11.
 GN CST11.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Hamil K.G., Hall S.H.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF501290; AAM21709.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 139 AA; 1686 MW; E1E36DB786B4D08C CRC64;
 QY Query Match 59.4%; Score 151; DB 11; Length 139;
 Best Local Similarity 58.1%; Pred. No. 1.5e-12;
 Matches 25; Conservative 12; Mismatches 6; Indels 0; Gaps 0;
 Db 1 NKESDDKYHFRIFRLVKVQROVTDHLEYHNVEMQWTTCKR 43
 56 NKSEDLNFRVVDILKSOEQITDSLEYEVNIAFTMCKR 96
 RESULT 3
 Q9DAP1 PRELIMINARY; PRT; 141 AA.
 AC Q9DAP1;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE 1700006C19R1K protein.
 GN 1700006C19R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; Tissue=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arai K., Hara A., Fukunishi Y., Kono H., Maehi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schiraldi L.M., Staudt F., Suzuki K., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Botfield D., Boujarda N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli S., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wyszewski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690(2001).
 RL EMBL; AK005665; BAB24175.1; -
 DR HSSP; P01038; ICEW.
 DR MGD; MGI:191654; 1700006C19R1K.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8B1AC378C CRC64;

Query Match 44.5%; Score 113; DB 11; Length 141;
 Best Local Similarity 51.2%; Pred. No. 2.1e-07;
 Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
 Db 1 NKESDDKYHFRIFRLVKVQROVTDHLEYHNVEMQWTTCKR 41
 56 NKASNDLNFVRVVDILKSOEQITDSLEYEVNIAFTMCKR 96
 RESULT 4
 Q80ZNS PRELIMINARY; PRT; 141 AA.
 AC Q80ZNS;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE RIKEN cDNA 1700006C19 gene.
 DE Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Testicle;
 RA Strauberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC048681; AAH48681.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR InterPro; IPR003243; Cystatin_C/M.
 DR Pfam; PF00031; Cystatin; 1.
 DR ProDom; PD001231; Cystatin_C/M; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8A84951F CRC64;
 QY Query Match 44.5%; Score 113; DB 11; Length 141;
 Best Local Similarity 51.2%; Pred. No. 2.1e-07;
 Matches 21; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
 Db 1 NKESDDKYHFRIFRLVKVQROVTDHLEYHNVEMQWTTCKR 41
 56 NKASNDLNFVRVVDILKSOEQITDSLEYEVNIAFTMCKR 96
 RESULT 5
 Q8WXUS PRELIMINARY; PRT; 103 AA.
 AC Q8WXUS;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE SC13delta.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.;
 RT "SC13, A novel epidermal specific member of the cystatin family."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF35481; AAL71992.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 SQ SEQUENCE 103 AA; 12285 MW; 05DD92C47387B022 CRC64;
 QY Query Match 43.3%; Score 110; DB 4; Length 103;
 Best Local Similarity 100.0%; Pred. No. 3.9e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 NKESDDKYHFRIFRLVKVQRO 21

Db 56 NKESDDKYHFRIFVLRVLRQ 76

RESULT 6

Q9EPX9 PRELIMINARY; PRT: 140 AA.

AC Q9EPX9 01-MAR-2001 (TRENBLREL. 16, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Cystatin C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=21010502; PubMed=11144350;
 RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakansson K., Grubb A.,
 Gage F.H.;
 RT "GF-2-Responsive neural stem cell proliferation requires Csg, a novel
 RT autocrine/paracrine cofactor.";
 RL Neuron 28:385-397(2000).
 DR EMBL; AF311741; AAC40283.1; -
 DR HSSP; P01034; 1G96.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.

FT CHAIN 21 140 CYSTATIN C.
 FT VARIANT 16 16 A -> G.
 FT VARIANT 84 84 L -> F.
 SQ SEQUENCE 140 AA; 15517 MW; 3A563406D58D785 CRC64;

Query Match 39.0%; Score 99; DB 11; Length 140;
 Best Local Similarity 43.2%; Pred. No. 1.6e-05;
 Matches 19; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

Qy 1 NKESDDKYHFRIFVLRVLRQVTDHLEHNLVEMQWTTCK 44
 Db 55 NKSNDAVHRSRAIOVAPAKQVAGVVFDDVEMGRITTSQT 98

RESULT 7

Q96SR4 PRELIMINARY; PRT: 112 AA.

AC Q96SR4 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Cystatin (Fragment).
 OS Acipenser sinensis (Chinese sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 NCBI_TaxID=61970;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two
 RT species of sturgeons.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF334610; AAK16731.1; -
 DR HSSP; P01038; 1A90.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.

FT NON TER 1 1
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFD8A08C00 CRC64;

Query Match 36.2%; Score 92; DB 13; Length 112;
 Best Local Similarity 43.9%; Pred. No. 0.00011;
 Matches 18; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy 1 NKESDDKYHFRIFVLRVLRQVTDHLEHNLVEMQWTTCK 41
 Db 28 NKASNDVYHRSKVRVLRQVAGVIGVYTVQWGRTSCK 68

RESULT 8

Q96SR3 PRELIMINARY; PRT: 112 AA.

AC Q96SR3 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Cystatin (Fragment).
 OS Acipenser schrenckii (Amur sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 NCBI_TaxID=11304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two
 RT species of sturgeons.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF334611; AAK16732.1; -
 DR HSSP; P01038; 1A90.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.

FT NON TER 1 1
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFD8A08C00 CRC64;

Query Match 36.2%; Score 92; DB 13; Length 112;
 Best Local Similarity 43.9%; Pred. No. 0.00011;
 Matches 18; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy 1 NKESDDKYHFRIFVLRVLRQVTDHLEHNLVEMQWTTCK 41
 Db 28 NKASNDVYHRSKVRVLRQVAGVIGVYTVQWGRTSCK 68

RESULT 9

Q9TY65 PRELIMINARY; PRT: 109 AA.

AC Q9TY65 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Antigen maltose binding protein (Fragment).
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94336252; PubMed=8058358;
 RA Trenholme K.R., Tree T.I., Gillespie A.J., Guiderian R., Matzele R.M.,
 Bradley J.B.;
 RT "Heterogeneity of IgG antibody responses to cloned Onchocerca volvulus
 RT antigens in microfiladermia positive individuals from Bameraiadas
 RT Province, Ecuador.";
 RL Parasite Immunol. 16:201-209(1994).
 DR EMBL; S71364; AAC60509.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.

```

DR InterPro; IPR000010; Cystatin.
DR Pfam; PF000031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT NON_TER 1
SQ SEQUENCE 109 AA; 12701 MW; 366B5EF14D8082BB CRC64;

Query Match
Best Local Similarity 34.1%; Score 81; DB 5; Length 109;
Matches 14; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 1 NKESDDKHFRIFRVLYKQROYTDHLEHNLVEMQWTTCK 41
Db 24 NQGSNDYHLMPIKLVKSSQVAGVKKYKMDVQVARSCK 64

RESULT 10
Q9NH95 PRELIMINARY; PRT; 148 AA.
AC Q9NH95;
DT 01-OCT-2000 (TReMBLrel. 15; Created)
DT 01-OCT-2000 (TReMBLrel. 15; Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24; Last annotation update)
DE Lb-cystatin.
OS Litomosoides sigmodontis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Litomosoides.
OX NCBI_Taxid=42156;
RN [1]
RP SEQUENCE FROM N.A.
RA Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;
RT "Characterization and immunological properties of a cystein protease
RT inhibitor of the filarial parasite Litomosoides sigmodontis.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229173; AAF35896.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
FT CHAIN 25 148
SQ SEQUENCE 148 AA; 16686 MW; 2350AA89CA5339C9 CRC64;

Query Match
Best Local Similarity 34.1%; Score 81; DB 5; Length 148;
Matches 14; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 1 NKESDDKHFRIFRVLYKQROYTDHLEHNLVEMQWTTCK 41
Db 62 NQGSNDYHLMPIKLVKSSQVAGVKKYKMDVQVARSCK 102

RESULT 11
O16159 PRELIMINARY; PRT; 161 AA.
ID O16159;
AC O16159;
DT 01-JAN-1998 (TReMBLrel. 05; Created)
DT 01-JAN-1998 (TReMBLrel. 05; Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24; Last annotation update)
DE Cystatin-type cysteine proteinase inhibitor.
DE BM-CPI-2.
OS Brugia malayi (Filarial nematode worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_Taxid=6279;
RN [1]
RP SEQUENCE FROM N.A.
RA Gregory W.F., Blaxter M.L., Maizels R.M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gregory W.F., Maizels R.M.;
RL "Two distinct cystatin-type cysteine protease inhibitors from the

```

```

RT Parasitic nematode Brugia malayi.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF015263; AAB6957.1; -
DR EMBL; AF177193; AAD51086.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
SQ SEQUENCE 161 AA; 18406 MW; 8081351EBE226EB5 CRC64;

Query Match
Best Local Similarity 34.1%; Score 80; DB 5; Length 161;
Matches 14; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 NKESDDKHFRIFRVLYKQROYTDHLEHNLVEMQWTTCK 41
Db 73 NQGSNDYHLMPIKLVKSSQVAGVKKYKMDVQVARSCK 113

RESULT 12
O80Y72 PRELIMINARY; PRT; 140 AA.
ID O80Y72;
AC O80Y72;
DT 01-JUN-2003 (TReMBLrel. 24; Created)
DT 01-JUN-2003 (TReMBLrel. 24; Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25; Last annotation update)
DE Cystatin-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22386257; Pubmed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marisla K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.S., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strauberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048646; A448646.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; CY; 1.
SQ SEQUENCE 140 AA; 16199 MW; 32633E99C4697DA0 CRC64;

Query Match
Best Local Similarity 30.3%; Score 77; DB 11; Length 140;
Matches 15; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

```

Qy 1 NKSDDKYNFRIFRVLKVRQVTDHLEHYHNVEMQWTTCCOK 44
Db 56 NNASNDTYLVQVQKLIQSQMQLTGVEXLVTVKIGRTCKKNET 99

RESULT 13

ID 012700 PRELIMINARY; PRT; 425 AA.
AC 012700;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Exo-1,3-beta-glucanase/1,3-beta-D-glucan glucanohydrolase
DE (EC 3.2.1.58).
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=27300;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 26077;
RX MEDLINE=99154256; PubMed=10029988;
RA Estebean P., del Rey F., Vazquez De Aldana C.R.;
RT "Cloning and characterization of 1,3-beta-glucanase-encoding genes
RT from non-conventional yeasts."
RL Yeast 15:91-109(1999).
DR EMBL; Z46871; CAA86951.1; -.
DR HSSP; P29711; 1CZ1.
DR GO; GO:0004338; F:glucan 1,3-beta-glucosidase activity; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . , IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001547; Glyco_hydro_5.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
KM GLYCOSIDASE; Hydrolase.
SQ SEQUENCE 425 AA; 49127 MW; 57F063ABE2FBF274 CRC64;
Query Match 29.9%; Score 76; DB 3; Length 425;
Best Local Similarity 28.8%; Pred. No. 0.066;
Matches 15; Conservative 10; Mismatches 11; Indels 16; Gaps 2;
Qy 6 DKYHFRIFRVLKVRQVTDHLEHYHNVEMQWTTCCOK 41
Db 277 DHHYVQFSGELQRSIDHITVACMGWDAKKYHNVAVAGWSAALTDCAR 328

RESULT 14

Q25620 PRELIMINARY; PRT; 125 AA.
AC 025620;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Beta-galactosidase.
GN LACZ.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92042729; PubMed=1840605;
RA Chandraashekar R., Masood K., Alvarez R.M., Ogunrinade A.F., Lujan R.,
RT "Molecular cloning and characterization of recombinant parasite
RT antigens for immunodiagnosis of onchocerciasis."
RL J. Clin. Invest. 88:1460-1466(1991).
DR EMBL; M60279; AAA29418.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.

SQ SEQUENCE 125 AA; 14369 MW; C73D5D9C7495A3CB CRC64;
Query Match 28.3%; Score 72; DB 5; Length 125;
Best Local Similarity 31.7%; Pred. No. 0.065;
Matches 13; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

Qy 1 NKSDDKYNFRIFRVLKVRQVTDHLEHYHNVEMQWTTCCOK 41
Db 40 NEQSDNENHLMPIKTLKSSQVAVGKXKMDVQVANSQCK 80

RESULT 15

ID 08K397 PRELIMINARY; PRT; 146 AA.
AC 08K397;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE RIKEN cDNA 110017E11 gene (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strauberg R.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027680; AAH27680.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 1.
FT NON TER 1
SQ SEQUENCE 146 AA; 16380 MW; 9D77BB9A063A5C4 CRC64;
Query Match 28.3%; Score 72; DB 11; Length 146;
Best Local Similarity 31.7%; Pred. No. 0.076;
Matches 13; Conservative 13; Mismatches 15; Indels 0; Gaps 0;
Qy 1 NKSDDKYNFRIFRVLKVRQVTDHLEHYHNVEMQWTTCCOK 41
Db 57 NMGSDSLYFRDTRKVIDAKQVLVAGIKYVLTLDIESIECK 97

Search completed: March 18, 2004, 14:21:28
Job time : 27.5635 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 21.7619 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-11
Perfect score: 138
Sequence: 1 RQVTDHLEHYHLNEMQWTTCKPE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	24	AAU79861	Aau79861 Human cys
2	138	100.0	46	AAU79860	Aau79860 Human cys
3	138	100.0	49	AAU79863	Aau79863 Human cys
4	138	100.0	52	AAU79864	Aau79864 Human cys
5	138	100.0	59	AAU79866	Aau79866 Human cys
6	138	100.0	80	AAU79865	Aau79865 Human cys
7	138	100.0	115	AAU79853	Aau79853 Human cys
8	138	100.0	117	AAU79854	Aau79854 Human cys
9	138	100.0	137	AAU79852	Aau79852 Human cys
10	133	96.4	33	AAU79862	Aau79862 Human cys
11	77	55.8	48	AAU79867	Aau79867 Human cys
12	71	51.4	142	ADD46708	Add46708 Rat Prote
13	71	51.4	142	ADD46704	Add46704 Rat Prote
14	63	45.7	142	AAE02404	Aae02404 Murine cy
15	63	45.7	142	AAE04433	Aae04433 Mouse cys
16	63	45.7	143	ADA14374	Ada14374 Mouse spe
17	60	43.5	141	AAV96576	Aav96576 Murine cy
18	60	43.5	141	AAE02403	Aae02403 Murine cy
19	60	43.5	141	AAE04432	Aae04432 Mouse tes
20	60	43.5	203	AAU83156	Aau83156 Novel sec
21	56	40.6	92	AAW78259	Aaw78259 Fragment
22	56	40.6	113	ADA57563	Ada57563 Human sec
23	56	40.6	113	ADA41457	Ada41457 Human sec
24	56	40.6	113	ADC74577	Adc74577 Human sec
25	56	40.6	113	ADD38088	Add38088 Human sec

ALIGNMENTS

RESULT 1
AAU79861 standard; peptide: 24 AA.
ID AAU79861;
AC AAU79861;
XX
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcy8) antigenic fragment #3.
XX
XX Cystatin-8; Zcy8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic peptide.
XX
XX Homo sapiens.
OS
PN W0200220567-A2.
XX
PD 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US026668.
XX
XX 01-SEP-2000; 2000US-0230230P.
XX
XX (ZYMO) ZYMOGENETICS INC.
PI Holloway JL, Gao Z, Bishop PD;
PI WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 97; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcy8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis; modulating seminal
CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcy8 is useful as research reagent for characterizing sites
CC of interaction between Zcy8 and its receptor. Zcy8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcy8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (1) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX

Sequence 24 AA;

Query Match 100.0%; Score 138; DB 5; Length 24;
 Best Local Similarity 100.0%; Pred. No. 3.3e-14;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHNLNVMQWTTCKKPE 24
 Db 1 RQVTDHLEHNLNVMQWTTCKKPE 24

RESULT 2

AAU79860 standard; peptide; 46 AA.

AAU79860;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #8.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting

spermatogenesis, and inhibiting cancer procoagulant protein which leads

to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)

polypeptide (1). (1) is useful for: inhibiting cancer procoagulant

protein in an individual and thus inhibiting the thrombotic events

associated with cancer; promoting spermatogenesis; modulating seminal

fluid viscosity; enhancing viability of cryopreserved sperm; sperm

motility and fertilisation; and as antigenic peptides to generate

antibodies. Zcys8 is useful as research reagent for characterising sites

of interaction between Zcys8 and its receptor. Zcys8 is useful in

enhancing fertilisation during assisted reproduction in humans and in

animals. Anti-(1) antibodies are useful to screen biological samples like

blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

presence of Zcys8. The antibodies are also useful to isolate large

quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

The polynucleotide encoding (1) is useful to detect and to localise the

expression of a Zcys8 gene in a biological sample and Zcys8

oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX

Sequence 46 AA;

Query Match 100.0%; Score 138; DB 5; Length 46;
 Best Local Similarity 100.0%; Pred. No. 7.2e-14;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHNLNVMQWTTCKKPE 24
 Db 20 RQVTDHLEHNLNVMQWTTCKKPE 43

RESULT 3

AAU79863 standard; peptide; 49 AA.

AAU79863;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #11.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;

spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

WO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting

spermatogenesis, and inhibiting cancer procoagulant protein which leads

to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97-98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)

polypeptide (1). (1) is useful for: inhibiting cancer procoagulant

protein in an individual and thus inhibiting the thrombotic events

associated with cancer; promoting spermatogenesis; modulating seminal

fluid viscosity; enhancing viability of cryopreserved sperm; sperm

motility and fertilisation; and as antigenic peptides to generate

antibodies. Zcys8 is useful as research reagent for characterising sites

of interaction between Zcys8 and its receptor. Zcys8 is useful in

enhancing fertilisation during assisted reproduction in humans and in

animals. Anti-(1) antibodies are useful to screen biological samples like

blood, urine, saliva, tissue biopsy and autopsy material in vitro for the

presence of Zcys8. The antibodies are also useful to isolate large

quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

The polynucleotide encoding (1) is useful to detect and to localise the

expression of a Zcys8 gene in a biological sample and Zcys8

oligonucleotide probes are useful for in vivo diagnosis. The

polynucleotide encoding (1) is useful in determining whether a subject's

chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene

copy number changes, insertions, deletions, restriction site changes and

rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 49 AA;

Query Match 100.0%; Score 138; DB 5; Length 49;
 Best Local Similarity 100.0%; Pred. No. 7,7e-14;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPE 24
 |||
 DB 23 RQVTDHLEHYHNVEMQWTTCKPE 46

RESULT 4
 AAU79864
 ID AAU79864 standard; peptide; 52 AA.
 XX
 AC AAU79864;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #12.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX W0200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis; modulating seminal
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX Sequence 52 AA;

Query Match 100.0%; Score 138; DB 5; Length 52;
 Best Local Similarity 100.0%; Pred. No. 8.3e-14;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTTCKPE 24
 |||
 DB 19 RQVTDHLEHYHNVEMQWTTCKPE 42

RESULT 5
 AAU79866
 ID AAU79866 standard; peptide; 59 AA.
 XX
 AC AAU79866;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #14.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX W0200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 99; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis; modulating seminal
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX Sequence 59 AA;

Query Match 100.0%; Score 138; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 9.6e-14;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
 ID 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
 DB 1 RQVTDHLEHYHLNVEMQWTTCKPE 24

RESULT 6
 AAU79865 standard; peptide; 80 AA.

AAU79865;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #13.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 spermatogenesis, and inhibiting cancer procoagulant protein which leads
 to inhibition of thrombotic events associated with cancer.

Claim 2; Page 98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
 polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 protein in an individual and thus inhibiting the thrombotic events
 associated with cancer; promoting spermatogenesis, modulating seminal
 fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 motility and fertilisation; and as antigenic peptides to generate
 antibodies. Zcys8 is useful as research reagent for characterising sites
 of interaction between Zcys8 and its receptor. Zcys8 is useful in
 enhancing fertilisation during assisted reproduction in humans and in
 animals. Anti-(I) antibodies are useful to screen biological samples like
 blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 presence of Zcys8. The antibodies are also useful to isolate large
 quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 The polynucleotide encoding (I) is useful to detect and to localise the
 expression of a Zcys8 gene in a biological sample and Zcys8
 oligonucleotide probes are useful for in vivo diagnosis. The
 polynucleotide encoding (I) is useful in determining whether a subject's
 chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 copy number changes, insertions, deletions, restriction site changes and
 rearrangements and genetic alterations that inactivate the Zcys8 gene.
 This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 80 AA;

Query Match 100.0%; Score 138; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.4e-13;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
 ID 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
 DB 22 RQVTDHLEHYHLNVEMQWTTCKPE 45

RESULT 7
 AAU79853 standard; protein; 115 AA.

AAU79853;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #1.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 sperm motility; fertilisation; antigenic fragment.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 spermatogenesis, and inhibiting cancer procoagulant protein which leads
 to inhibition of thrombotic events associated with cancer.

Claim 2; Page 94; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
 polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 protein in an individual and thus inhibiting the thrombotic events
 associated with cancer; promoting spermatogenesis, modulating seminal
 fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 motility and fertilisation; and as antigenic peptides to generate
 antibodies. Zcys8 is useful as research reagent for characterising sites
 of interaction between Zcys8 and its receptor. Zcys8 is useful in
 enhancing fertilisation during assisted reproduction in humans and in
 animals. Anti-(I) antibodies are useful to screen biological samples like
 blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 presence of Zcys8. The antibodies are also useful to isolate large
 quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 The polynucleotide encoding (I) is useful to detect and to localise the
 expression of a Zcys8 gene in a biological sample and Zcys8
 oligonucleotide probes are useful for in vivo diagnosis. The
 polynucleotide encoding (I) is useful in determining whether a subject's
 chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 copy number changes, insertions, deletions, restriction site changes and
 rearrangements and genetic alterations that inactivate the Zcys8 gene.
 This sequence represents an antigenic fragment of human cystatin-8
 (Zcys8)

Sequence 115 AA;

Query Match 100.0%; Score 138; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
 ID 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
 DB 52 RQVTDHLEHYHLNVEMQWTTCKPE 75

RESULT 8
 AAU79854

```

ID  AAU79854 standard; protein; 117 AA.
XX
AC  AAU79854;
XX
DT  15-JUL-2002 (first entry)
XX
DE  Human cystatin-8 (Zcys8) antigenic fragment #2.
XX
KM  Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM  spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KM  sperm motility; fertilisation; antigenic fragment.
XX
OS  Homo sapiens.
XX
PN  WO200220567-A2.
XX
PD  14-MAR-2002.
XX
PF  29-AUG-2001; 2001WO-US026868.
XX
PR  01-SEP-2000; 2000US-0230230P.
XX
PA  (ZYMO ) ZYMOGENETICS INC.
XX
PI  Holloway JL, Gao Z, Bishop PD;
XX
DR  WPI; 2002-383044/41.
XX
PT  Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT  spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT  to inhibition of thrombotic events associated with cancer.
XX
PS  Claim 2; Page 94-95; 100pp; English.
XX
CC  The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC  polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC  protein in an individual and thus inhibiting the thrombotic events
CC  associated with cancer; promoting spermatogenesis, modulating seminal
CC  fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC  motility and fertilisation; and as antigenic peptides to generate
CC  antibodies. Zcys8 is useful as research reagent for characterising sites
CC  of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC  enhancing fertilisation during assisted reproduction in humans and in
CC  animals. Anti-(I) antibodies are useful to screen biological samples like
CC  blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC  presence of Zcys8. The antibodies are also useful to isolate large
CC  quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC  The polynucleotide encoding (I) is useful to detect and to localise the
CC  expression of a Zcys8 gene in a biological sample and Zcys8
CC  oligonucleotide probes are useful for in vivo diagnosis. The
CC  polynucleotide encoding (I) is useful in determining whether a subject's
CC  chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC  copy number changes, insertions, deletions, restriction site changes and
CC  rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC  This sequence represents an antigenic fragment of human cystatin-8
CC  (Zcys8)
XX
SQ  Sequence 117 AA;
XX
Query Match      100.0%; Score 138; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 ROVTDHLEHYHIANVEMQWTTCKRPE 24
Db  54 ROVTDHLEHYHIANVEMQWTTCKRPE 77
XX
RESULT 9
AAU79852
ID  AAU79852 standard; protein; 137 AA.
XX
AC  AAU79852;

```

```

XX
DT  15-JUL-2002 (first entry)
XX
DE  Human cystatin-8 (Zcys8) .
XX
KM  Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM  spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KM  sperm motility; fertilisation.
XX
OS  Homo sapiens.
XX
PN  WO200220567-A2.
XX
PD  14-MAR-2002.
XX
PF  29-AUG-2001; 2001WO-US026868.
XX
PR  01-SEP-2000; 2000US-0230230P.
XX
PA  (ZYMO ) ZYMOGENETICS INC.
XX
PI  Holloway JL, Gao Z, Bishop PD;
XX
DR  WPI; 2002-383044/41.
XX
DR  N-PSDB; ABK49522.
XX
PT  Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT  spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT  to inhibition of thrombotic events associated with cancer.
XX
PS  Claim 2; Page 93-94; 100pp; English.
XX
CC  The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC  polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC  protein in an individual and thus inhibiting the thrombotic events
CC  associated with cancer; promoting spermatogenesis, modulating seminal
CC  fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC  motility and fertilisation; and as antigenic peptides to generate
CC  antibodies. Zcys8 is useful as research reagent for characterising sites
CC  of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC  enhancing fertilisation during assisted reproduction in humans and in
CC  animals. Anti-(I) antibodies are useful to screen biological samples like
CC  blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC  presence of Zcys8. The antibodies are also useful to isolate large
CC  quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC  The polynucleotide encoding (I) is useful to detect and to localise the
CC  expression of a Zcys8 gene in a biological sample and Zcys8
CC  oligonucleotide probes are useful for in vivo diagnosis. The
CC  polynucleotide encoding (I) is useful in determining whether a subject's
CC  chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC  copy number changes, insertions, deletions, restriction site changes and
CC  rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC  This is the amino acid sequence of human cystatin-8 (Zcys8)
XX
SQ  Sequence 137 AA;
XX
Query Match      100.0%; Score 138; DB 5; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.6e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 ROVTDHLEHYHIANVEMQWTTCKRPE 24
Db  74 ROVTDHLEHYHIANVEMQWTTCKRPE 97
XX
RESULT 10
AAU79862
ID  AAU79862 standard; peptide; 33 AA.
XX
AC  AAU79862;
XX
DT  15-JUL-2002 (first entry)
XX

```

DE Human cystatin-8 (Zcys8) antigenic fragment #10.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
XX sperm motility; fertilisation; antigenic peptide.
XX
XX Homo sapiens.
OS
XX
XX WO200220567-A2.
XX
XX 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US026866.
XX
XX 01-SEP-2000; 2000US-0230230P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX
XX Holloway JL, Gao Z, Bishop PD;
XX
XX WPI: 2002-383044/41.
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
XX
XX Claim 2, Page 97, 100pp; English.
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(1) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (1) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vitro diagnosis. The
CC polynucleotide encoding (1) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
XX Sequence 33 AA;
SQ
XX
XX Query Match 96.4%; Score 133; DB 5; Length 33;
XX Best Local Similarity 100.0%; Pred. No. 2,9e-13;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX 2 QVTDHLEHYHLNVEMQWTCCKPE 24
XX ||||||||||||||||||||
XX 1 QVTDHLEHYHLNVEMQWTCCKPE 23
XX
XX
XX RESULT 11
XX AAU79867
XX ID AAU79867 standard; peptide; 48 AA.
XX
XX AAU79867;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human cystatin-8 (Zcys8) antigenic fragment #15.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
XX spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
XX

```

XX sperm motility; fertilisation; antigenic peptide.
OS Homo sapiens.
XX
XX WO200220567-A2.
XX
XX 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US026686.
XX
XX 01-SEP-2000; 2000US-0230230P.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Holloway JL, Gao Z, Bishop PD;
XX
XX WPI; 2002-383044/41.
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
XX spermatogenesis, and inhibiting cancer procoagulant protein which leads
XX to inhibition of thrombotic events associated with cancer.
XX
XX
XX Claim 2; Page 99; 100pp; English.
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
XX polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
XX protein in an individual and thus inhibiting the thrombotic events
XX associated with cancer; promoting spermatogenesis, modulating seminal
XX fluid viscosity, enhancing viability of cryopreserved sperm, sperm
XX motility and fertilisation; and as antigenic peptides to generate
XX antibodies. Zcys8 is useful as research reagent for characterising sites
XX of interaction between Zcys8 and its receptor. Zcys8 is useful in
XX enhancing fertilisation during assisted reproduction in humans and in
XX animals. Anti-(I) antibodies are useful to screen biological samples like
XX blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
XX presence of Zcys8. The antibodies are also useful to isolate large
XX quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
XX The polynucleotide encoding (I) is useful to detect and to localise the
XX expression of a Zcys8 gene in a biological sample and Zcys8
XX oligonucleotide probes are useful for in vivo diagnosis. The
XX polynucleotide encoding (I) is useful in determining whether a subject's
XX chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
XX copy number changes, insertions, deletions, restriction site changes and
XX rearrangements and genetic alterations that inactivate the Zcys8 gene.
XX This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
XX
XX Sequence 48 AA;
XX
XX Query March 55.8%; Score 77; DB 5; Length 48;
XX Best Local Similarity 100.0%; Pred. NO. 0.00021;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 12 NVEMQMTTCCKPE 24
XX |||||
XX 1 NVEMQMTTCCKPE 13
XX
XX
XX RESULT 12
XX ID ADD46708
XX AC ADD46708 standard; protein; 142 AA.
XX
XX ADD46708;
XX
XX 29-JAN-2004 (first entry)
XX
XX Rat Protein AAC6317, SEQ ID NO 12393.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.

```

XX 27-FEB-2003.
 PD 14-AUG-2002; 2002MO-US025765.
 PF 14-AUG-2001; 2001US-0312147P.
 XX 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; AAC6317.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 142 AA;
 Query Match 51.4%; Score 71; DB 7; Length 142;
 Best Local Similarity 45.5%; Pred. No. 0.0064;
 Matches 10; Conservative 8; Mismatches 4; Indels 0; Gaps 0;
 QY 2 QVTDHLEHYHNVEMQWTCCKP 23
 Db 77 QITDRMEYHIDVOISRNCCKP 98
 RESULT 13
 ADD6704 standard; protein; 142 AA.
 AC ADD6704;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein AAC6317, SEQ ID NO 12389.
 XX
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.
 OS
 XX WO2003016475-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002MO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; AAC6317.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 142 AA;
 Query Match 51.4%; Score 71; DB 7; Length 142;
 Best Local Similarity 45.5%; Pred. No. 0.0064;
 Matches 10; Conservative 8; Mismatches 4; Indels 0; Gaps 0;
 QY 2 QVTDHLEHYHNVEMQWTCCKP 23
 Db 77 QITDRMEYHIDVOISRNCCKP 98
 RESULT 14
 AAE02404 standard; protein; 142 AA.
 AC AAE02404;
 XX
 DT 10-AUG-2001 (first entry)
 XX

DE Murine cystatin-related epididymal specific protein (CRES).
 XX
 XX Murine; Cystatin T; zcy83; cystatin-related epididymal specific protein;
 KM CRES; inhibitor; cysteine proteinase; male reproductive tissue; testis;
 KM spermatogenesis; therapy; reproductive disorder.
 XX
 OS Mus musculus.
 XX
 PN US6235708-B1.
 XX
 PD 22-MAY-2001.
 XX
 PF 01-NOV-1999; 99US-00431480.
 XX
 PR 20-NOV-1998; 98US-0109217P.
 PR 28-SEP-1999; 99US-0156382P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Feldhaus AL;
 XX
 DR WPI; 2001-342846/36.
 XX
 CC Cystatin T polypeptides are useful for modulating spermatogenesis and
 PT studying, diagnosing and treating reproductive disorders.
 XX
 PS Disclosure; Col 45-46; 32pp; English.
 XX
 CC The present invention relates to cystatin T (also known as zcy83) DNA and
 CC protein sequences. Cystatin T is testis specific and is homologous to
 CC cystatin-related epididymal specific gene (CRES) and type 2 cystatins.
 CC Cystatins inhibit cysteine proteinases and are found with male
 CC reproductive tissues and secretions. Cystatin T sequence is useful for
 CC modulating spermatogenesis and studying, diagnosing and treating
 CC reproductive disorders. The present sequence is murine cystatin-related
 CC epididymal specific (CRES) protein
 XX
 SQ Sequence 142 AA;
 XX
 QY 2 QVTDLEHYHNVEMQWTTCKP 23
 Db 77 QITDRMEYQIDVQISRSNCKKP 98
 XX
 Query Match 45.7%; Score 63; DB 4; Length 142;
 Best Local Similarity 40.9%; Pred. No. 0.11;
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 XX
 RESULT 15
 AAE04433
 ID AAE04433 standard; protein; 142 AA.
 XX
 AC AAE04433;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Mouse cystatin-related epididymal specific (CRES) protein.
 XX
 KM Mouse; cystatin T; zcy83; testis specific; spermatogenesis modulator;
 KM cystatin-related epididymal specific gene; CRES; gene therapy;
 KM sperm production; antifertility.
 XX
 OS Mus musculus.
 XX
 PN US6245529-B1.
 XX
 PD 12-JUN-2001.
 XX
 PF 17-JUL-2000; 2000US-00617302.
 XX
 PR 20-NOV-1998; 98US-0109217P.
 PR 28-SEP-1999; 99US-0156382P.
 PR 01-NOV-1999; 99US-00431480.

XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Feldhaus AL;
 XX
 DR WPI; 2001-407271/43.
 XX
 CC New polynucleotides encoding testis-specific cystatin-like protein
 PT cystatin T, useful in gene therapy for modulating cystatin T activity,
 PT particularly for modulating spermatogenesis, or enhancing sperm
 PT production or fertility.
 XX
 PS Disclosure; Col 47-48; 33pp; English.
 XX
 CC The present sequence is mouse cystatin-related epididymal specific (CRES)
 CC protein which is homologous to mouse testis specific cystatin T (also
 CC known as zcy83). The cystatin T polynucleotide is useful in gene therapy
 CC applications, where it is desired to increase or inhibit cystatin T
 CC activity. It is also useful for producing cystatin T polypeptide, as well
 CC as for detecting the expression of a cystatin T gene in a biological
 CC sample. The cystatin T is useful for modulating spermatogenesis, and may
 CC be used to study or modulate that function in in vitro or in vivo
 CC systems. In particular, it is also useful for enhancing sperm production,
 CC increasing the number of viable sperm in a sample, or enhancing
 CC fertilisation
 XX
 SQ Sequence 142 AA;
 XX
 QY 2 QVTDLEHYHNVEMQWTTCKP 23
 Db 77 QITDRMEYQIDVQISRSNCKKP 98
 XX
 Search completed: March 18, 2004, 14:15:00
 Job time : 22.7619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 6.14286 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-11

Perfect score: 138
Sequence: 1 RQVTDHLEHYHLNEMQWTTCKPK 24Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patente AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCtus_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	45.7	142	3	US-09-431-480-3
2	63	45.7	142	3	US-09-617-302-3
3	60	43.5	141	3	US-09-431-480-2
4	60	43.5	141	3	US-09-617-302-2
5	56	40.6	142	3	US-09-431-480-4
6	56	40.6	142	3	US-09-617-302-4
7	53	38.4	127	4	US-08-849-303-19
8	52	37.7	112	4	US-08-849-303-16
9	52	37.7	118	4	US-09-775-932-24
10	50	36.2	68	4	US-08-858-207A-492
11	50	36.2	714	2	US-08-472-534-3
12	49	35.5	140	4	US-09-886-319A-46
13	49	35.5	140	4	US-09-886-319A-48
14	49	35.5	145	2	US-08-832-535-2
15	49	35.5	145	3	US-09-019-485-2
16	49	35.5	145	3	US-09-019-485-3
17	49	35.5	145	3	US-09-431-480-9
18	49	35.5	145	3	US-09-617-302-9
19	49	35.5	145	4	US-09-528-436B-2
20	49	35.5	178	2	US-08-791-522-1
21	49	35.5	178	3	US-09-314-777-13
22	48.5	34.8	111	4	US-08-849-303-26
23	48	34.8	116	4	US-09-775-932-16
24	48	34.8	120	6	5432264-4
25	48	34.8	120	6	5432264-4
26	48	34.8	120	6	5432264-4
27	48	34.8	139	2	US-08-791-522-4

28	48	34.8	139	3	US-09-314-777-4	Sequence 4, Appli
29	48	34.8	139	4	US-08-849-303-15	Sequence 15, Appl
30	48	34.8	145	2	US-08-832-535-11	Sequence 11, Appl
31	48	34.8	146	2	US-08-791-522-3	Sequence 3, Appli
32	48	34.8	146	3	US-08-744-138-3	Sequence 4, Appli
33	48	34.8	146	3	US-09-019-485-4	Sequence 3, Appli
34	48	34.8	146	3	US-09-314-777-3	Sequence 3, Appli
35	48	34.8	146	3	US-09-431-480-6	Sequence 6, Appli
36	48	34.8	146	3	US-09-617-302-6	Sequence 6, Appli
37	48	34.8	146	4	US-09-241-376-3	Sequence 3, Appli
38	48	34.8	146	4	US-09-528-436B-3	Sequence 3, Appli
39	48	34.8	146	4	US-09-886-319A-47	Sequence 47, Appl
40	48	34.8	146	4	US-09-940-497-3	Sequence 3, Appli
41	48	34.8	146	4	US-09-976-594-17	Sequence 37, Appl
42	48	34.8	146	4	US-08-849-303-17	Sequence 17, Appl
43	48	34.8	146	5	PCT-US95-07135-9	Sequence 9, Appli
44	48	34.8	146	6	5432264-6	Patent No. 5432264
45	48	34.8	443	4	US-09-107-532A-7119	Sequence 7119, Ap

ALIGNMENTS

```
RESULT 1
US-09-431-480-3
; Sequence 3, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; PILE REFERENCE: 98-72
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-480-3

Query Match          45.7%; Score 63; DB 3; Length 142;
Best Local Similarity 40.9%; Pred. No. 0.0092;
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY      2 QVTDHLEHYHLNEMQWTTCKPK 23
       |||:||||:|:|:|:|:|:|
Db      77 QITRMEYQIDVOISRSNCKPK 98

RESULT 2
US-09-617-302-3
; Sequence 3, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; PILE REFERENCE: 98-72 C1
; CURRENT FILING DATE: 2000-07-17
; EARLIER APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 142
TYPE: PRT
ORGANISM: Mus musculus
US-09-617-302-3

Query Match 45.7%; Score 63; DB 3; Length 142;
Best Local Similarity 40.9%; Pred. No. 0.0092;
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTTCKP 23
DB 77 QITDRMEYQIDVQISRSNCKP 98

RESULT 3
US-09-431-480-2
Sequence 2, Application US/09431480
Patent No. 6235708
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72
CURRENT APPLICATION NUMBER: US/09/431,480
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 60/109,217
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/156,382
EARLIER FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-431-480-2

Query Match 43.5%; Score 60; DB 3; Length 141;
Best Local Similarity 52.4%; Pred. No. 0.028;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTTCKP 22
DB 76 QITDSELYLEVNIARTWCKK 96

RESULT 4
US-09-617-302-2
Sequence 2, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-617-302-2

Query Match 43.5%; Score 60; DB 3; Length 141;
Best Local Similarity 52.4%; Pred. No. 0.028;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTTCKP 22
DB 76 QITDSELYLEVNIARTWCKK 96

RESULT 5
US-09-431-480-4
Sequence 4, Application US/09431480
Patent No. 6235708
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72
CURRENT APPLICATION NUMBER: US/09/431,480
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 60/109,217
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/156,382
EARLIER FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 142
TYPE: PRT
ORGANISM: Homo sapiens
US-09-431-480-4

Query Match 40.6%; Score 56; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 0.13;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTTCKP 23
DB 77 QVTNLELYLIDVEIARSDCRKP 98

RESULT 6
US-09-617-302-4
Sequence 4, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 142
TYPE: PRT
ORGANISM: Homo sapiens
US-09-617-302-4

Query Match 40.6%; Score 56; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 0.13;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTTCKP 23
DB 77 QVTNLELYLIDVEIARSDCRKP 98

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentln Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/849,303
      FILING DATE: 21-MAY-1997
      CLASSIFICATION: 514
      ATTORNEY/AGENT INFORMATION:
      NAME: Jackson Esq., David A.
      REGISTRATION NUMBER: 26,742
      REFERENCE/DOCKET NUMBER: 1321-1-003
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: 201-487-5800
      TELEFAX: 201-343-1684
      TELEX: 133521
      INFORMATION FOR SEQ ID NO: 16:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 112 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      HYPOTHEICAL: NO
US-08-849-303-16

Query Match          37.7%; Score 52; DB 4; Length 112;
Best Local Similarity 41.7%; Pred. No. 0.44;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 RQVTDHLEHYHLNVEMOWTTCOKPE 24
       :|::||::||::||::||::||::||:
DB      47 KQVSGMNYFLDVELGRITCTKSQ 70

RESULT 9
US-09-775-932-24
; Sequence 24, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cysteatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 24
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-775-932-24

Query Match          37.7%; Score 52; DB 4; Length 118;
Best Local Similarity 41.7%; Pred. No. 0.47;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 RQVTDHLEHYHLNVEMOWTTCOKPE 24
       :|::||::||::||::||::||::||:
DB      53 KQVSGMNYFLDVELGRITCTKSQ 76

RESULT 10
US-08-858-207A-492
; Sequence 492, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John

```

APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: No. 6348328e1 Compounds
NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,207A
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017670
FILING DATE: 14-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50475
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 492:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6348328e
US-08-858-207A-492

Query Match 36.2%; Score 50; DB 4; Length 68;
Best Local Similarity 37.9%; Pred. No. 0.51;
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Qy 1 RQVTDH-----LEYHLNVEMQWTTCKP 23
Db 30 RQWTDHPNGDFMETPLNTQFDWNGIRKP 58

RESULT 11
US-08-472-534-3
Sequence 3, Application US/08472534
Patent No. 5919620
GENERAL INFORMATION:
APPLICANT: Hamel, Josee
APPLICANT: Brodeur, Bernard R
APPLICANT: Martin, Denis
TITLE OF INVENTION: HEAT SHOCK PROTEIN HSP72 FROM
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,534
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr, James F
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: Biovac-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 714 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-534-3

Query Match 36.2%; Score 50; DB 2; Length 714;
Best Local Similarity 37.9%; Pred. No. 9.1;
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Qy 1 RQVTDH-----LEYHLNVEMQWTTCKP 23
Db 301 RQWTDHPNGDFMETPLNTQFDWNGIRKP 329

RESULT 12
US-09-886-319A-46
Sequence 46, Application US/09886319A
Patent No. 6586185
GENERAL INFORMATION:
APPLICANT: Wolf, Eckard
APPLICANT: Werner, Sabine
APPLICANT: Haller, Jorn-Peter
APPLICANT: Regenbogen, Johannes
APPLICANT: Goppelt, Andreas
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
FILE REFERENCE: 50125/014002
CURRENT APPLICATION NUMBER: US/09/886,319A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 60/222,081
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: DE 10030149.5
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 140
TYPE: PRT
ORGANISM: Mus musculus
US-09-886-319A-46

Query Match 35.5%; Score 49; DB 4; Length 140;
Best Local Similarity 41.7%; Pred. No. 1.8;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RQVTDHLEYHLNVEMQWTTCKP 24
Db 74 KQLVAGVYFLDVEWGRITCTKSG 97

RESULT 13
US-09-886-319A-48
Sequence 48, Application US/09886319A
Patent No. 6586185
GENERAL INFORMATION:
APPLICANT: Wolf, Eckard

APPLICANT: Werner, Sabine
APPLICANT: Halle, Jorn-Peter
APPLICANT: Regendogen, Johannes
APPLICANT: Goppelt, Andreas
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
FILE REFERENCE: 50125/014002
CURRENT APPLICATION NUMBER: US/09/886,319A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 60/222,081
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: DE 10030149.5
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 48
LENGTH: 140
TYPE: PRT
ORGANISM: Mus musculus
US-09-886-319A-48

Query Match 35.5%; Score 49; DB 4; Length 140;
Best Local Similarity 41.7%; Pred. No. 1.8;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 QVTDHLEHYLNVEMQWTTCKPE 24
DB 74 KQLVAGVNYPLDVEIMGRTCTCKSQ 97

RESULT 14
US-08-832-535-2
Sequence 2, Application US/08832535
Patent No. 5919658
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: LI, HAODONG
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER J
TITLE OF INVENTION: HUMAN CYSTATIN F
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832,535
FILING DATE: 03-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KIMBALL, PAUL C.
REGISTRATION NUMBER: 34,610
REFERENCE/DOCKET NUMBER: PP265
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1744
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-832-535-2

Query Match 35.5%; Score 49; DB 2; Length 145;
Best Local Similarity 47.6%; Pred. No. 1.9;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 QVTDHLEHYLNVEMQWTTCK 22
DB 81 QIVGKLKYMVEIGRTTCK 101

RESULT 15
US-09-019-485-2
Sequence 2, Application US/09019485
Patent No. 6066617
GENERAL INFORMATION:
APPLICANT: LI, Haodong
APPLICANT: YU, Guo-Liang
APPLICANT: Gentz, Reiner
APPLICANT: Ni, Jian
TITLE OF INVENTION: Cystatin F
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,485
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Robert H.
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PP265P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3013098439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-485-2

Query Match 35.5%; Score 49; DB 3; Length 145;
Best Local Similarity 47.6%; Pred. No. 1.9;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 QVTDHLEHYLNVEMQWTTCK 22
DB 81 QIVGKLKYMVEIGRTTCK 101

Search completed: March 18, 2004, 14:25:36
Job time: 6.14286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 14.5238 Seconds

(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-11

Perfect score: 138
Sequence: 1 RQVTDHLEHYHIANVEMQWTTCKPE 24Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppaa/US09A_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	24	9	US-09-941-314-11 Sequence 11, Appl
2	138	100.0	46	9	US-09-941-314-10 Sequence 10, Appl
3	138	100.0	49	9	US-09-941-314-13 Sequence 13, Appl
4	138	100.0	52	9	US-09-941-314-14 Sequence 14, Appl
5	138	100.0	59	9	US-09-941-314-16 Sequence 16, Appl
6	138	100.0	80	9	US-09-941-314-15 Sequence 15, Appl
7	138	100.0	115	9	US-09-941-314-3 Sequence 3, Appl
8	138	100.0	117	9	US-09-941-314-4 Sequence 4, Appl
9	138	100.0	137	9	US-09-941-314-2 Sequence 2, Appl
10	133	96.4	33	9	US-09-941-314-12 Sequence 12, Appl
11	77	55.8	48	9	US-09-941-314-17 Sequence 17, Appl
12	60	43.5	103	9	US-09-893-737-164 Sequence 164, Appl
13	56	40.6	138	14	US-10-239-663-37 Sequence 37, Appl
14	55	39.9	37	9	US-09-864-761-40966 Sequence 40966, A
15	55	39.9	145	9	US-09-740-638-2 Sequence 2, Appl

16	55	39.9	145	13	US-10-006-467-2	Sequence 2, Appl
17	55	39.9	145	14	US-10-235-148-2	Sequence 2, Appl
18	55	39.9	155	9	US-09-740-638-5	Sequence 5, Appl
19	55	39.9	165	13	US-10-006-467-5	Sequence 5, Appl
20	55	39.9	165	14	US-10-235-148-5	Sequence 16, Appl
21	53	38.4	127	8	US-08-849-303-19	Sequence 19, Appl
22	52	37.7	112	8	US-08-849-303-16	Sequence 16, Appl
23	52	37.7	118	9	US-09-775-932-24	Sequence 24, Appl
24	51.5	37.3	678	12	US-10-282-122A-61129	Sequence 61129, A
25	50	36.2	145	14	US-10-168-425-14	Sequence 14, Appl
26	49	35.5	140	14	US-10-376-564-46	Sequence 46, Appl
27	49	35.5	140	14	US-10-376-564-48	Sequence 48, Appl
28	49	35.5	145	14	US-10-329-428-2	Sequence 2, Appl
29	49	35.5	167	10	US-09-746-783-197	Sequence 197, App
30	49	35.5	178	9	US-09-969-834-1	Sequence 1, Appl
31	48	34.8	111	8	US-08-849-303-26	Sequence 26, Appl
32	48	34.8	116	9	US-09-775-932-16	Sequence 16, Appl
33	48	34.8	120	9	US-09-775-932-2	Sequence 2, Appl
34	48	34.8	139	8	US-08-849-303-15	Sequence 15, Appl
35	48	34.8	139	9	US-09-969-834-4	Sequence 4, Appl
36	48	34.8	146	8	US-08-849-303-17	Sequence 17, Appl
37	48	34.8	146	9	US-09-940-497-3	Sequence 3, Appl
38	48	34.8	146	9	US-09-969-834-3	Sequence 3, Appl
39	48	34.8	146	14	US-10-329-428-3	Sequence 3, Appl
40	48	34.8	146	14	US-10-376-564-47	Sequence 47, Appl
41	48	34.8	438	12	US-10-282-122A-57752	Sequence 57752, A
42	47	34.1	273	14	US-10-156-761-12874	Sequence 12874, A
43	46	33.3	122	9	US-09-775-932-10	Sequence 10, Appl
44	46	33.3	142	8	US-08-849-303-20	Sequence 20, Appl
45	46	33.3	142	9	US-09-940-497-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-941-314-11
Sequence 11, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
PRIOR FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-11

Query Match 100.0%; Score 138; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHIANVEMQWTTCKPE 24
DB 1 RQVTDHLEHYHIANVEMQWTTCKPE 24

RESULT 2
US-09-941-314-10
Sequence 10, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

```

; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-10

```

```

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 46;
Pred. No. 3,2e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RQVTDHLEHYHNLNEMQWTTCKP 24
Db 20 RQVTDHLEHYHNLNEMQWTTCKP 43

```

```

RESULT 3
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: Zymogenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-13

```

```

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 49;
Pred. No. 3,4e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RQVTDHLEHYHNLNEMQWTTCKP 24
Db 23 RQVTDHLEHYHNLNEMQWTTCKP 46

```

```

RESULT 4
US-09-941-314-14
; Sequence 14, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: Zymogenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-14

```

```

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 52;
Pred. No. 3,7e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RQVTDHLEHYHNLNEMQWTTCKP 24
Db 19 RQVTDHLEHYHNLNEMQWTTCKP 42

```

```

RESULT 5
US-09-941-314-15
; Sequence 15, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: Zymogenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-16

```

```

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 59;
Pred. No. 4,2e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RQVTDHLEHYHNLNEMQWTTCKP 24
Db 1 RQVTDHLEHYHNLNEMQWTTCKP 24

```

```

RESULT 6
US-09-941-314-15
; Sequence 15, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: Zymogenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-15

```

```

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 80;
Pred. No. 5,7e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RQVTDHLEHYHNLNEMQWTTCKP 24
Db 22 RQVTDHLEHYHNLNEMQWTTCKP 45

```

```

RESULT 7
US-09-941-314-3
; Sequence 3, Application US/09941314
; Patent No. US20020142396A1

```



```

; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-3

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 115;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
Db 52 RQVTDHLEHYHLNVEMQWTTCKPE 75

RESULT 8
US-09-941-314-4
; Sequence 4, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-4

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 117;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
Db 54 RQVTDHLEHYHLNVEMQWTTCKPE 77

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-2

Query Match
Best Local Similarity 100.0%; Score 138; DB 9; Length 137;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHLNVEMQWTTCKPE 24
Db 74 RQVTDHLEHYHLNVEMQWTTCKPE 97

RESULT 10
US-09-941-314-12
; Sequence 12, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-12

Query Match
Best Local Similarity 96.4%; Score 133; DB 9; Length 133;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QVTDHLEHYHLNVEMQWTTCKPE 24
Db 1 QVTDHLEHYHLNVEMQWTTCKPE 23

RESULT 11
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941.314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-17

Query Match
Best Local Similarity 55.8%; Score 77; DB 9; Length 48;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NVEMQWTTCKPE 24
Db 1 NVEMQWTTCKPE 13

```

```
RESULT 12
US-09-893-737-164
; Sequence 164, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 164
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-164

Query Match      43.5%; Score 60; DB 9; Length 203;
Best Local Similarity 52.4%; Pred. No. 0.61;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      2 QVTDHLEYHLNVEWQMTTCOK 22
        ||| |||:: ||:: ||: ||
Db      115 QVTDSELYEYIEVKIARTICK 135

RESULT 13
US-10-239-663-37
; Sequence 37, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kohnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: G950018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-37

Query Match      40.6%; Score 56; DB 14; Length 138;
Best Local Similarity 47.6%; Pred. No. 1.6;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40986
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121894.14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EST HUMAN HIT: AM665128.1, EVALUATE 1.10e-02
; OTHER INFORMATION: SWISSPROT HIT: O88969, EVALUATE 1.00e-04
US-09-864-761-40986

Query Match      39.9%; Score 55; DB 9; Length 37;
Best Local Similarity 50.0%; Pred. No. 0.57;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      3 VTDHLEYHLNVEWQMTTCOK 22
        ||| |||:: ||:: ||: ||
Db      1 VTDSELYEYIEVKIARTICK 20
```

```

RESULT 15
US-09-740-638--2
; Sequence 2, Application US/09740638
; Patent No. US20020006656A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Zcy45: A Member of the Cystatatin
; TITLE OF INVENTION: Superfamily
; FILE REFERENCE: 99-104
; CURRENT APPLICATION NUMBER: US/09/740,638
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-740-638--2

```

Query Match	39.9%	Score 55;	DB 9;	Length 145;
Best Local Similarity	40.9%;	Pred. No. 2.4;		
Matches	9;	Conservative	7;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
Qy      1 RQVTDHLEYHLNVEMQWTTCK 22
          ||::||:|::|||::
Db      72 RQLTGVEYIVTKIGWTKCKR 93
```

Search completed: March 18, 2004, 14:30:54
Job time : 14.5238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 4.61905 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-11
Perfect score: 138
Sequence: 1 RQVTDHLEHYHLNEMQWTTQCKPE 24

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	45.7	139	2 A45361	cystatin-related e
2	53	38.4	120	2 S10587	cystatin C - rat
3	53	38.4	127	2 S07085	cystatin C precurs
4	53	38.4	1779	2 T23130	hypothetical prote
5	52	37.7	112	1 UD80	cystatin - bovine
6	50	36.2	232	2 T28369	probable uracil-DN
7	50	36.2	588	2 C95252	L-fucose isomerase
8	48	34.8	111	2 A99717	cystatin - puff ad
9	48	34.8	139	1 UDCH	cystatin precursor
10	48	34.8	146	1 UDHU	cystatin C precurs
11	48	34.8	519	2 B88322	protein C47D12.8 l
12	47.5	34.4	519	2 T11998	hypothetical prote
13	47.5	34.4	1308	2 B75198	DNA helicase relat
14	47.5	34.4	456	2 B84960	adenylosuccinate l
15	47	34.1	555	2 S39953	isocitrate lyase (
16	47	34.1	371	2 B96614	hypothetical prote
17	46.5	33.7	133	2 JC4536	cystatin precursor
18	46	33.3	142	2 A47142	cystatin D precurs
19	46	33.3	355	2 T34405	hypothetical prote
20	46	33.3	677	1 SYECWT	histidinol-phospha
21	46	33.3	677	1 SYECWT	methionine-CRNA 1i
22	46	33.3	677	2 C85839	methionine-CRNA 1i
23	46	33.3	677	2 H90993	methionine-CRNA 1i
24	46	33.3	677	2 H90993	methionine-CRNA 1i
25	45.5	33.0	652	2 S58666	serine/threonine-p
26	45.5	33.0	1640	2 H88094	protein F399.2 (i
27	45	32.6	140	2 A36163	cystatin C precurs
28	45	32.6	454	2 T27249	hypothetical prote
29	45	32.6	455	2 T00856	pectate lyase (EC

30	45	32.6	482	2 T15333	hypothetical prote
31	45	32.6	675	2 AF0185	methionine-CRNA 1i
32	45	32.6	913	2 G64110	hypothetical prote
33	44.5	32.2	2055	2 T31617	hypothetical prote
34	44	31.9	142	2 AC1642	hypothetical prote
35	44	31.9	436	1 KGBOL1	kininogen, LMW 1 p
36	44	31.9	447	2 AC1279	kininogen, LMW 1 p
37	44	31.9	447	2 D71812	udp-n-acetylmuram
38	44	31.9	497	2 A84956	kininogen, HMW 1 p
39	44	31.9	621	1 KGBOL1	protein F1E22.16 l
40	44	31.9	1075	2 C96682	DNA-directed RNA p
41	44	31.9	1478	2 S78131	hypothetical prote
42	43.5	31.5	419	2 T25577	protein K03H1.2 (i
43	43.5	31.5	1131	2 F88570	hypothetical prote
44	43.5	31.5	1451	2 S41025	fiber - human aden
45	43	31.2	325	2 D37476	

ALIGNMENTS

RESULT 1
A45361
cystatin-related epididymal specific protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C/Accession: A45361
R/Cornwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.
Mol. Endocrinol. 6, 1653-1664, 1992
A/Title: The CREB gene: a unique testis-regulated gene related to the cystatin family is
A/Reference number: A45361; MUID:93078799; PMID:1280328
A/Accession: A45361
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-139 <COR>
A/Cross-references: GB:949926; MID:9260492; PIDN:AAC35390.1; PID:9260493
A/Note: sequence extracted from NCBI backbone (NCBI:118813)
C/Superfamily: cystatin; cystatin homology
F/28-139/Domain: cystatin homology <CYS>

Query Match 45.7%; Score 63; DB 2; Length 139;
Best Local Similarity 40.9%; Pred. No. 0.029;
Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QVTDHLEHYHLNEMQWTTQCKP 23
Db 74 QVTDHLEHYHLNEMQWTTQCKP 95

RESULT 2
S10587
cystatin C - rat
C/Species: Rattus sp. (rat)
C/Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C/Accession: S10587
R/Barnard, F.; Barnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Brillard, M.; Gauth
Biochem. Biophys. Res. Commun. 161, 166, 1990
A/Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylat
A/Reference number: S10587; MUID:90380276; PMID:2400577
A/Accession: S10587
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-120 <ESN>
A/Note: 43-Asn was also found
A/Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala
C/Superfamily: cystatin; cystatin homology
F/9-120/Domain: cystatin homology <CYS>

Query Match 38.4%; Score 53; DB 2; Length 120;
Best Local Similarity 41.7%; Pred. No. 0.81;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 RQVTDHLEHYHLNEMQWTTQCKP 24

Db 54 KQVAGINYYLDVEMGRITCKSQ 77

RESULT 3

S07085

Cysteatin C precursor - rat (fragment)

C/Species: Rattus norvegicus (Norway rat)

C/Date: 01-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999

C/Accession: S07085; S01337; S21109

R/Ennard, T.; Dickson, P.W.; Ennard, F.; Averill, S.; Ribbridger, G.P.; Gauthier, F.; Sch

Eur. J. Biochem. 186, 35-42, 1989

A/Title: The cDNA structure and expression analysis of the genes for the cysteine prote

A/Reference number: S07085; MUID:90092122; PMID:2689174

A/Accession: S07085

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-127 <COL>

A/Cross-references: EMBL:X16957; NID:G56041; PIDN:CAA34831.1; PID:G736290

R/Ennard, A.; Ennard, F.; Faucher, D.; Gauthier, F.

FEBS Lett. 236, 475-478, 1988

A/Title: Two rat homologues of human cystatin C.

A/Reference number: S01337; MUID:88313020; PMID:3044831

A/Accession: S01337

A/Molecule type: protein

A/Residues: 8-49 <ESN>

R/Ennard, A.; Ennard, F.; Gulliou, F.; Gauthier, F.

FEBS Lett. 300, 131-135, 1992

A/Title: Production of the cysteine proteinase inhibitor cystatin C by rat Sertoli cells

A/Reference number: S21109; MUID:9225121; PMID:1565513

A/Accession: S21109

A/Molecule type: protein

A/Residues: 8; 'XX', 11-20 <ES2>

A/Superfamily: cystatin; cystatin homology

C/Keywords: cysteine proteinase inhibitor

F/16-127/Domain: cystatin homology <CTS>

F/80-90, 104-124/Disulfide bonds: #status predicted

Query Match 38.4%; Score 53; DB 2; Length 127;

Best Local Similarity 41.7%; Pred. No. 0.86;

Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 RQVTDHLEVHNVEMQWTCQKPE 24

Db 61 KQVAGINYYLDVEMGRITCKSQ 84

RESULT 4

T23130

hypothetical protein T28B8.4 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000

C/Accession: T23130; T23403

R/White, S.

Submitted to the EMBL Data Library, December 1997

A/Reference number: Z19690

A/Accession: T23130

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1779 <WIL>

A/Cross-references: EMBL:AL021066; PIDN:CAA15925.1; GSPDB:GNO0019; CESP:T28B8.4

A/Experimental source: clone H31B20

R/White, S.

Submitted to the EMBL Data Library, October 1996

A/Reference number: Z20029

A/Accession: T23403

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1779 <W12>

A/Cross-references: EMBL:Z81133; PIDN:CAB03445.1; GSPDB:GNO0019; CESP:T28B8.4

A/Experimental source: clone T28B8

C/Genetics:

A/Genes: CESP:T28B8.4

A/Map position: 1
A/Intons: 161/2; 223/2; 309/3; 332/2; 547/3; 603/1; 657/3; 745/2; 802/1; 856/1; 1031/1;
C/Superfamily: Caenorhabditis elegans hypothetical protein T28B8.3

Query Match 38.4%; Score 53; DB 2; Length 1779;
Best Local Similarity 37.0%; Pred. No. 15;
Matches 10; Conservative 5; Mismatches 4; Indels 8; Gaps 1;

QY 1 RQVTDHLEVHNVEMQWTCQKPE 24

Db 749 RRTDHDYIINNPLTKHPTNQMWT 775

RESULT 5

UDBO

Cysteatin - bovine

N/Alternate names: thiol proteinase inhibitor

C/Species: Bos primigenius taurus (cattle)

C/Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 06-Dec-1996

C/Accession: A01271

R/Hirado, M.; Tsunawawa, S.; Sakiyama, F.; Ninobe, M.; Fujii, S.

FEBS Lett. 186, 41-45, 1985

A/Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase int

A/Reference number: A01271; MUID:85231205; PMID:3891407

A/Accession: A01271

A/Molecule type: protein

A/Residues: 1-112 <HIR>

C/Superfamily: cystatin; cystatin homology

C/Keywords: colostrum; cysteine proteinase inhibitor

F/2-112/Domain: cystatin homology <CTS>

F/48-52/Region: inhibitory #status predicted

F/66-76, 90-110/Disulfide bonds: #status predicted

Query Match 37.7%; Score 52; DB 1; Length 112;

Best Local Similarity 41.7%; Pred. No. 1.1;

Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 RQVTDHLEVHNVEMQWTCQKPE 24

Db 47 KQVAGINYYLDVEMGRITCKSQ 70

RESULT 6

T28369

probable uracil-DNA glycosylase (EC 3.2.2.-) UNG MSV208 [similarity] - Melanoplus sanguin

C/Species: Melanoplus sanguinipes entomopoxvirus

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C/Accession: T28369

R/Alonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.

J. Virol. 73, 533-552, 1999

A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.

A/Reference number: Z20484; MUID:99102612; PMID:9847359

A/Accession: T28369

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-232 <AFO>

A/Cross-references: EMBL:AF063866; NID:G4049647; PIDN:AAC97753.1; PID:G4049793

C/Genetics:

A/Note: MSV208

C/Superfamily: uracil-DNA glycosylase

C/Keywords: glycosidase; hydrolase

Query Match 36.2%; Score 50; DB 2; Length 232;

Best Local Similarity 31.6%; Pred. No. 4.7;

Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEVHNVEMQWTCQKPE 20

Db 212 KANEYLKIKKNIETIDWISC 230

RESULT 7

C95252

L-fucose isomerase [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: C95252
 R:Retellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95252
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-588 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK76212.1; PID:G14973669; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP2158
 C:Superfamily: isomerase fucI

Query Match 36.2%; Score 50; DB 2; Length 588;
 Best Local Similarity 37.9%; Pred. No. 13;
 Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 1 RQVTDH-----LEVHLNVEKQMTTCCKP 23
 Db 301 RQWTDHPNGDFMETPLNTQDWMGIRKP 329

RESULT 8
 A:Species: Streptococcus pneumoniae (strain R6)
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: A99717
 R:Hoskins, J.A.; Albhorn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: A99717
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-588 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAU00766.1; PID:G15459665; GSPDB:GN00174
 C:Genetics:
 A:Gene: fucI
 C:Superfamily: isomerase fucI
 C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 36.2%; Score 50; DB 2; Length 588;
 Best Local Similarity 37.9%; Pred. No. 13;
 Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 1 RQVTDH-----LEVHLNVEKQMTTCCKP 23
 Db 301 RQWTDHPNGDFMETPLNTQDWMGIRKP 329

RESULT 9
 C:Species: Bifidobacterium (puff adder)
 C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 30-Sep-1993
 C:Accession: A28793
 R:Rittonja, A.; Evans, H.J.; Machleidt, W.; Barrett, A.J.
 Biochem. J. 246, 799-802, 1987
 A:Title: Amino acid sequence of a cystatin from venom of the African puff adder (Bifidob
 A:Reference number: A28793; MUID:88076861; PMID:3500714
 A:Accession: A28793

A:Molecule type: protein
 A:Residues: 1-111 <RT>
 C:Superfamily: cystatin; cystatin homology

Query Match 34.8%; Score 48; DB 2; Length 111;
 Best Local Similarity 42.9%; Pred. No. 4.2;
 Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEVHLNVEKQMTTCCKP 22
 Db 47 QVSVGVKXYLMMEBLKTTCKK 67

RESULT 10
 UDCH
 Cystatin precursor - chicken
 N:Alternate names: cystatin 1; cysteine proteinase inhibitor; egg-white cystatin
 C:Species: Gallus gallus (chicken)
 C>Date: 03-Aug-1984 #sequence_revision 12-Apr-1996 #text_change 29-Oct-1999
 C:Accession: A34456; A01274; S01461; S48159; S04008; JN0789
 R:Coletta, R.; Sakaguchi, Y.; Nagase, H.; Bird, J.W.C.
 J. Biol. Chem. 264, 17164-17169, 1989
 A:Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue
 A:Reference number: A34456; MUID:9008873; PMID:2793849
 A:Accession: A34456
 A:Molecule type: mRNA
 A:Residues: 1-139 <COL>
 A:Cross-references: GB:J05077; NID:G211714; PIDN:AAA48744.1; PID:G211715
 R:Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.K.; Barrett, A.J.
 Biochem. J. 217, 813-817, 1984
 A:Title: Cystatin. Amino acid sequence and possible secondary structure.
 A:Reference number: A01274; MUID:84178305; PMID:6712597
 A:Accession: A01274
 A:Molecule type: protein
 A:Residues: 24-139 <SCH>
 R:Turk, V.; Brixin, J.; Longer, M.; Rittonja, A.; Eropkin, M.; Borchardt, U.; Machleidt, W.
 Hope-Seayler's Z. Physiol. Chem. 364, 1487-1496, 1983
 A:Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin
 A:Reference number: S01461; MUID:84110059; PMID:6662498
 A:Accession: S01461
 A:Molecule type: protein
 A:Residues: 24-139 <TUR>
 R:Anastasi, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.
 Biochem. J. 211, 129-138, 1983
 A:Title: Cystatin, a protein inhibitor of cysteine proteinases. Improved purification fr
 A:Reference number: A37514; MUID:83256421; PMID:6409085
 A:Contents: annotation; characterization of protein
 R:Grubb, A.; Lofberg, H.; Barrett, A.J.
 FEBS Lett. 170, 370-374, 1984
 A:Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.
 A:Reference number: S01462
 A:Contents: annotation; disulfide bonds
 R:Auerwald, E.A.; Naegler, D.K.; Schulze, A.J.; Engh, R.A.; Genger, G.; Machleidt, W.
 Eur. J. Biochem. 224, 407-415, 1994
 A:Title: Production, inhibitory activity, folding and conformational analysis of an N-ter
 A:Reference number: S48159; MUID:95010016; PMID:7925354
 A:Accession: S48159
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 24-139 <AUE>
 R:Reber, B.; Kriegelstein, K.; Hennechen, A.; Koe, J.; Turk, V.; Huber, R.; Bode, W.
 FEBS Lett. 248, 162-168, 1989
 A:Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.
 A:Reference number: S04008; MUID:89252033; PMID:2721673
 A:Accession: S04008
 A:Molecule type: protein
 A:Residues: 97-114 <LAB>
 R:Coletta, R.; Bird, J.W.C.
 Gene 130, 175-181, 1993
 A:Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping to
 A:Reference number: JN0789; MUID:93366172; PMID:8359584
 A:Accession: JN0789
 A:Molecule type: DNA

A;Residues: 1-139 <CO2>
 A;Cross-references: GB:M95725
 A;Note: authors failed to translate the codon for residue 115-Tyr
 C;Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases
 C;Genetics:
 A;Gene: Cpn
 A;Introns: 76/3; 114/3
 C;Superfamily: cystatin; cystatin homology
 C;Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-139/Product: cystatin, long form #status experimental <CYLF>
 F;30-139/Domain: cystatin homology <CYS>
 F;32-139/Product: cystatin, short form #status experimental <CYSF>
 F;76-80/Region: inhibitory #status predicted
 F;94-104, 118-138/Disulfide bonds: #status experimental
 F;103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 34.8%; Score 48; DB 1; Length 139;
 Best Local Similarity 45.5%; Pred. No. 5.4;
 Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Cy 1 RQVTDHEHNVEMQWTTCK 22
 Db 75 RQVSGIKYILQVEIRGTTCPK 96

RESULT 11
 UDHU
 cystatin C precursor [validated] - human
 N;Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma
 C;Species: Homo sapiens (man)
 C;Date: 06-Jul-1992 #sequence revision 31-Mar-1991 #text change 08-Dec-2000
 C;Accession: S10216; S000004; J10095; A33400; S02751; A01270; A25434; S12288; A32732; A60
 R;Abrahamson, M.; Olafsson, I.; Palsdottir, A.; Ulfvback, M.; Lundwall, A.; Jenson, O.
 Biochem. J. 268, 287-294, 1990
 A;Title: Structure and expression of the human cystatin C gene.
 A;Reference number: S10216; MUID:90303202; PMID:2363674
 A;Accession: S10216
 A;Molecule type: DNA
 A;Residues: 1-146 <AB1>
 A;Cross-references: EMBL:X52255; NID:G30257; PIDN:CAA36497.1; PID:G296643
 R;Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.
 FEBS Lett. 216, 229-233, 1987
 A;Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the
 A;Reference number: S00004; MUID:87219149; PMID:3495457
 A;Accession: S00004
 A;Molecule type: mRNA
 A;Residues: 1-146 <AB2>
 A;Cross-references: EMBL:X05607; NID:G30371; PIDN:CAA29096.1; PID:G755738
 R;Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.
 J. Exp. Med. 169, 1771-1778, 1989
 A;Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a
 A;Reference number: J10095; MUID:89235594; PMID:2541223
 A;Accession: J10095
 A;Molecule type: DNA
 A;Residues: 1-146 <LEW>
 A;Cross-references: GB:X61681; NID:G30367; PIDN:CAA43856.2; PID:G4490944
 A;Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit
 e)
 R;Salton, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada,
 Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989
 A;Title: The human cystatin C gene (C93) is a member of the cystatin gene family which
 A;Reference number: A33400; MUID:89350949; PMID:2764935
 A;Accession: A33400
 A;Molecule type: DNA
 A;Residues: 1-24, 'T', 26-146 <SA1>
 A;Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:G181385; PIDN:AAA52164.1; PID:G
 R;Ghiso, J.; Cowan, N.; Frangione, B.
 Biol. Chem. Hoppe-Seyler 369, 205-208, 1988
 A;Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron
 A;Reference number: S02751; MUID:89076507; PMID:3264504
 A;Accession: S02751
 A;Molecule type: DNA

A;Residues: 82-119 <GH2>
 A;Cross-references: EMBL:M27769
 A;Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown i
 R;Grubb, A.; Lofberg, H.
 Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982
 A;Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the
 A;Reference number: A01270; MUID:82222268; PMID:6283552
 A;Accession: A01270
 A;Molecule type: protein
 A;Residues: 27-131, 'S', 133-146 <GRU>
 R;Ghiso, J.; Jenson, O.; Frangione, B.
 Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986
 A;Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland ty
 A;Reference number: A25434; MUID:86206076; PMID:3517880
 A;Accession: A25434
 A;Molecule type: protein
 A;Residues: 37-93, 'Q', 95-146 <GHI>
 R;Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983
 A;Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatir
 A;Reference number: S01461; MUID:84110059; PMID:6662498
 A;Accession: S12288
 A;Molecule type: protein
 A;Residues: 27-73 <TUR>
 R;Brzin, J.; Popovic, T.; Turk, V.
 Biochem. Biophys. Res. Commun. 118, 103-109, 1984
 A;Title: Human cystatin, a new protein inhibitor of cysteine proteinases.
 A;Reference number: A32732; MUID:84128015; PMID:6365094
 A;Accession: A32732
 A;Molecule type: protein
 A;Residues: 27-76 <B2>
 R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jenson, O.; Grubb, A.
 Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
 A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyste
 A;Reference number: A60552; MUID:90193615; PMID:2315647
 A;Accession: A60552
 A;Molecule type: protein
 A;Residues: 27-49, 'XX', 52-64 <OLA>
 A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
 e defective gene is not present in CSF but is found instead in amyloid deposits
 R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
 Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
 A;Title: Different forms of human cystatin C.
 A;Reference number: S10607; MUID:91025625; PMID:2222856
 A;Accession: S10607
 A;Molecule type: protein
 A;Residues: 27-53 <POP>
 A;Experimental source: urine, kidney disease
 A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were als
 R;Grubb, A.; Lofberg, H.; Barrett, A.J.
 FEBS Lett. 170, 370-374, 1984
 A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.
 A;Reference number: S01462
 A;Contents: annotation; disulfide bonds
 R;Berti, P.J.; Storer, A.C.
 Biochem. J. 302, 411-416, 1994
 A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility
 A;Reference number: S55305; MUID:94379969; PMID:8092921
 A;Accession: S55305
 A;Molecule type: protein
 A;Status: preliminary
 A;Residues: 27-49, 106-146 <BER>
 C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl
 e patients with certain autoimmune diseases.
 C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an import
 C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorr
 C;Genetics:
 A;Gene: GDB:G973
 A;Cross-references: GDB:119817; OMIM:105150
 A;Map position: 20p11.2-20p11.2
 A;Introns: 81/3; 119/3
 C;Superfamily: cystatin; cystatin homology
 C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline

F:1-26/Domain: signal sequence #status predicted <SIG>
E:27-146/Product: cystatin C #status experimental <MAT>
F:35-146/Domain: cystatin homology <CYS>
F:81-85/Region: inhibitory #status predicted
F:129/Modified site: hydroxyproline (Pro) (partial) #status experimental
F:199-109,123-143/Disulfide bonds: #status experimental

Query Match	34.8%	Score 48	DB 1	Length 146
Best Local Similarity	37.5%	Pred. No. 5.7		
Matches 9	Conservative 6	Mismatches 9	Indels 0	Gaps 0

```

QY      1 ROVTDHLEHYHLNVEMQWTTCCQKPE 24
      ::: | | | | | | | | | |
DB      80 KQIVAGVNVFLDVELGRITCTKIQ 103

```

```

RESULT 12
B88332
protein C47D12.8 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88332
R:Anonymous, the C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:95063613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Accession: B88332
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,519 <STO>
A:Cross-references: GB:chr_II; PIDN:CAA93766.1; PID:g3875010; GSFDB:GN00020; CESP:C47D12
C:Genetics:
A:Gene: C47D12.8
A:Map position: 2

```

Query Match	34.4%	Score 47.5;	DB 2;	Length 519;
Best Local Similarity	52.6%	Pred. No. 27;		
Matches	10;	Conservative	4;	Mismatches
			4;	Indels
				1;
				Gaps
				1;
Qy	1	ROVTDHLEYHNVEMQTT	19	
Db	408	ROVTDVVRKGIN-KKIKMT	425	

```

RESULT 13
T19998
hypothetical protein C47D12.8 - Caenorhabditis elegans (fragment)
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T19998
R/Gajadaty, S.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19209
A/Accession: T19998
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-519 <WIL>
A/Cross-references: EMBL:Z69902; PIDN:CA93766.1; GSPDB:GN00020; CESP:C47D12.8
C/Experimental source: clone C47D12
C/Genetics:
A/Gene: CESP:C47D12.8
A/Map position: 2
A/Introns: 147/2; 394/1

```

Query Match	34.4%	Score 47.5	DB 2	Length 519
Best Local Similarity	52.6%	Pred. No. 27		
Matches	10	Conservative	4	Mismatches 4; Indels 1; Gaps 1;
Qy	1	ROVTDHLEHYLANVEMQMTT	19	
			: :	
Db	408	ROVTDVVRGGIN-KKKMMT	425	

RESULT 14
B75198
DNA helicase related protein PAB0067 - *Pyrococcus abyssi* (strain Oreyay)
C.Species: *Pyrococcus abyssi*
C.Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C.Accession: B75198
R.anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A.Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure
A.Reference number: A75001
A.Accession: B75198
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1308 <KAM>
A.Cross-references: GB:AL096836; GB:AL096836; NID:G5457433; PIDN:CB49033.1; PID:el51492
A.Experimental source: strain Oreyay
C.Genetics:
A.Gene: PAB0067

```

QY      2 QVTDHLEHYHNVEMOW-----TTCQ 21
      :  :  :  :  :  :  :  :  :  :
DB      212 EFVDTIEKLNLEIKMSSEPIITACTTCQ 240

Query Match          34.4%; Score 47.5; DB 2; length 1308;
Best Local Similarity 34.5%; Pred. No. 73;
Matches 10; Conservative 5; Mismatches 5; Indels 9; Gaps 1;

```

```

RESULT 15
B84960
adenylosuccinate lyase (BC 4.3.2.2) [imported] - Buchnera sp. (strain APS)
C.Species: Buchnera sp.
C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C.Accession: B84960
R.Shigenobu, S.; Matanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
A.Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A.Reference number: A84930; MUID:20445173; PMID:10993077
A.Accession: B84960
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-456 <STO>
A.Cross-References: GB:AP000398; GSPDB:GN00144
A.Experimental source: strain APS
C.Genetics:
A.Gene: purB; BU263
C.Superfamily: fumarate hydratase
C.Keywords: amidine-lyase; carbon-nitrogen lyase

Query Match          34.1%; Score 47; DB 2; Length 456;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 LEYHLNVEQW 17
   | | | | |
DB 33 LKYLRLIEYQW 43

Search completed: March 18, 2004, 14:23:16
Job time : 5.61905 secs

```

Search completed: March 18, 2004, 14:23:16
Job time : 5.61905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 2.85714 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-11
Perfect score: 138
Sequence: 1 RQYTDHLEHTLHNMOWTCKRPE 24

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	137	CS11_HUMAN	Q9H112 homo sapien
2	91	65.9	139	CS11_MOUSE	O94269 mus musculu
3	71	51.4	142	CST8_RAT	O88969 rattus norv
4	63	45.7	142	CST8_MOUSE	P32766 mus musculu
5	56	40.6	142	CST8_HUMAN	O60676 homo sapien
6	55	39.9	165	CSTL_HUMAN	Q9H114 homo sapien
7	53	38.4	127	CYT8_RAT	P14841 rattus norv
8	52	37.7	148	CYT8_BOVIN	P01035 bos taurus
9	50	36.2	146	CYT8_SAISC	O19093 saimiri sci
10	49	35.5	140	CYT8_MOUSE	P24460 mus musculu
11	49	35.5	145	CYTF_HUMAN	O76096 homo sapien
12	48	34.8	111	CYT_BITAR	P08935 bitis ariet
13	48	34.8	139	CYT_CHICK	P01038 gallus galli
14	48	34.8	146	CYT8_HUMAN	P01034 homo sapien
15	48	34.8	146	CYT8_MOUSE	O97862 mus musculu
16	48	34.8	148	CYT8_RABIT	O97862 mus musculu
17	47	34.1	146	PUR8_BUCAT	P57351 buchnera ap
18	47	34.1	541	ACEA_YARLI	P41555 yarrowia li
19	46	33.3	142	CYT8_HUMAN	P08325 homo sapien
20	46	33.3	144	CYT8_MOUSE	O88098 mus musculu
21	46	33.3	367	HIS8_SYNY3	P73807 synecocyst
22	46	33.3	676	SYM_ECO57	O87677 escherichia
23	46	33.3	676	SYM_ECO16	O87677 escherichia
24	46	33.3	676	SYM_ECO11	O87677 escherichia
25	46	33.3	759	PEPX_STRGN	O93442 streptococc
26	46	33.3	1145	Y623_DROME	O94767 drosophila
27	45.5	33.0	652	SMU1_HUMAN	P59817 homo sapien
28	45.5	33.0	652	SSP1_SCHPO	P65526 schizosacch
29	45	32.6	455	PEL6_ARATH	O64510 arabidopsis
30	45	32.6	485	MURE_FUSNP	O86901 yersinia pe
31	45	32.6	675	SYM_YERP	O86901 yersinia pe
32	45	32.6	913	YCI17_HAEIN	P45114 haemophilus
33	44.5	32.2	107	TIM4_BOVIN	O97563 bos taurus

34	44.5	32.2	170	1	TIM4_RABIT	O97591 oryctolagus
35	44.5	32.2	224	1	TIM4_HUMAN	O99727 homo sapien
36	44.5	32.2	543	1	SWT2_HUMAN	O86972 homo sapien
37	44.5	32.2	1227	1	PR16_HUMAN	O92620 homo sapien
38	44	31.9	436	1	KNL1_BOVIN	P01046 bos taurus
39	44	31.9	447	1	MURE_HELPJ	O92366 helicobacte
40	44	31.9	497	1	MURE_BUCAT	P57316 buchnera ap
41	44	31.9	621	1	KNH1_BOVIN	P01044 bos taurus
42	44	31.9	3670	1	CSM3_HUMAN	O72407 homo sapien
43	43.5	31.5	1131	1	MOG1_CASEL	P34456 caenorhabdi
44	43	31.2	325	1	PIBP_ADELP	P35774 human adeno
45	43	31.2	459	1	PEL3_ARATH	O96962 arabidopsis

ALIGNMENTS

RESULT 1
ID CS11_HUMAN STANDARD; PRT; 137 AA.
AC Q9H112; Q9H113; 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin II precursor.
GN CS11 OR CST8L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOPFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.B., Cobby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leivasalho M.H., Leverkus M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurtry A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H., Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skene C.D., Smith M.L., Soderlund C., Steward C.A., Sultoni J.B., Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;
RL Nature 414:665-671(2001).
RT The DNA sequence and comparative analysis of human chromosome 20.;
RL Nature 414:665-671(2001).
CC -1- SUBCUTANEOUS LOCATION: Secreted (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9H112-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H112-2; Sequence=VSP_001260;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the cystatin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC EMBL; AL096677; CAC13170.1; -

CC EMBL; AL096677; CAC17423.1; -

CC HSSP; P01038; 1A90.

CC Genew; HGNC:15959; CST11.

CC InterPro: IPR000010; Cystatin.

CC Pfam; PF00031; cystatin. 1.

CC SMART; SM00043; CY; 1.

CC PROSITE; PS00287; CYSTATIN; FALSE NEG.

CC Thiol protease inhibitor; Signal; Alternative splicing.

CC SIGNAL 1 25 POTENTIAL.

CC CHAIN 26 137 CYSTATIN 11.

CC SITE 75 79 SECONDARY AREA OF CONTACT (POTENTIAL).

CC DISULFID 93 101 BY SIMILARITY.

CC DISULFID 114 134 BY SIMILARITY.

CC CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).

CC VARSPIC 76 110 Missing (in isoform 2).

CC SEQUENCE 137 AA; 16375 MW; C585C8C39A585C3B CRC64;

Query Match 100.0%; Score 138; DB 1; Length 137;

Best Local Similarity 100.0%; Pred. No. 6.5e-14;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHILNVEMQWTTCKPE 24

DB 74 RQVTDHLEHILNVEMQWTTCKPE 97

RESULT 2

CS11_MOUSE STANDARD; PRT; 139 AA.

ID CS11_MOUSE

AC Q9D269;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cystatin 11 precursor.

GN CST11.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Epididymis;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I., Saito T., Okazaki Y., Gotohori T., Bono H., Kanukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schirral L.M., Seubert F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Botfield D., Boujarda N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gusticich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H., Lyons P.J., Marchionni L., Mashima Y., Mazzarelli J., Mombereis P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakanato N., Sasakawa H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F., Suzuki H., Togo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilting L., Wyshak-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;

RA "Functional annotation of a full-length mouse cDNA collection.";

RT Nature 409:685-690(2001).

CC -1- SUBCELLULAR LOCATION: Secreted (Potential).

CC -1- SIMILARITY: Belongs to the cystatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC EMBL; AK020300; BAB32061.1; -

CC HSSP; P01034; 1G96.

CC MED; MGI:1925490; Cst11.

CC InterPro: IPR000010; Cystatin.

CC Pfam; PF00031; cystatin. 1.

CC SMART; SM00043; CY; 1.

CC PROSITE; PS00287; CYSTATIN; FALSE NEG.

CC Thiol protease inhibitor; Signal.

CC SIGNAL 1 28 POTENTIAL.

CC CHAIN 29 139 CYSTATIN 11.

CC SITE 76 80 SECONDARY AREA OF CONTACT (POTENTIAL).

CC DISULFID 94 102 BY SIMILARITY.

CC DISULFID 115 135 BY SIMILARITY.

CC CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).

CC SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;

Query Match 65.9%; Score 91; DB 1; Length 139;

Best Local Similarity 70.8%; Pred. No. 8.5e-07;

Matches 17; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RQVTDHLEHILNVEMQWTTCKPE 24

DB 75 RQVTDHLEHILNVEMQWTTCKPE 98

RESULT 3

CST8_RAT STANDARD; PRT; 142 AA.

ID CST8_RAT

AC O88969;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin B).

GN CST8 OR CRES.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;

RX MEDLINE=99247899; PubMed=10229662;

RA Cornwall G.A., Hsia N., Sutton H.G.;

RT "Structure, alternative splicing and chromosomal localization of the Cystatin-related epididymal spermatogenic gene.";

RL Biochem. J. 340:85-93(1999).

CC -1- FUNCTION: Performs a specialized role during sperm development and maturation.

CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC EMBL; AF090692; AAC6317.1; -

CC HSSP; P01034; 1G96.

CC InterPro: IPR000010; Cystatin.

CC Pfam; PF00031; cystatin. 1.

CC SMART; SM00043; CY; 1.

KW Thiol protease inhibitor; Signal.
 PT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 142 CYPSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT SITE 77 81 PROTEIN.
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 119 139 BY SIMILARITY.
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 142 AA; 16246 MW; PB873FPA6B6CAB34 CRC64;
 Query Match 51.4%; Score 71; DB 1; Length 142;
 Best Local Similarity 45.5%; Pred. No. 0.00092;
 Matches 10; Conservative 8; Mismatches 4; Indels 0; Gaps 0;
 QY 2 QVTDHLEHYLVEMQWTCCKP 23
 77 QITDMEHYIDVOISRNCCKP 98
 Db
 RESULT 4
 CST8 MOUSE STANDARD; PRT; 142 AA.
 ID CST8_MOUSE
 AC P32766; 089102;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-
 related epididymal specific protein) (Cystatin 8).
 GN CST8 OR CRGS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H, and CD-1;
 RX MEDLINE=99247899; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.,
 RT "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermatogenic gene.";
 RL Biochem. J. 340:85-93(1999).
 RN [2]
 RP SEQUENCE OF 4-142 FROM N.A.
 RC TISSUE=Epididymis;
 RX MEDLINE=93078799; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.,
 RT "The CRGS gene: a unique testis-regulated gene related to the cystatin
 RT family is highly restricted in its expression to the proximal region
 RT of the mouse epididymis.";
 RL Mol. Endocrinol. 6:1653-1664(1992).
 RN [1]
 RP FUNCTION: Performs a specialized role during sperm development and
 RP maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively spermatised by the proximal caput epididymium.
 CC -1- INDUCTION: Testicular factors or hormones other than androgens
 CC present in the testicular fluid may be involved in the regulation
 CC of CRGS gene expression.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF091503; AAC61754.1; -
 CC DR EMBL; AF090691; AAC36316.1; -
 CC EMBL; S49926; AAC35390.1; -

DR PIR: A45361; A45361.
 DR HSSP: P01034; 1G96.
 DR MGD; MGI:107161; Gsc8.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 KW Thiol protease inhibitor; Signal.
 PT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 142 CYPSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT SITE 77 81 PROTEIN.
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 119 139 BY SIMILARITY.
 FT CARBOHYD 33 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 4 15 PMLSLFLFTIP -> GTRDQVGSQK (IN REF. 2).
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;
 Query Match 45.7%; Score 63; DB 1; Length 142;
 Best Local Similarity 40.9%; Pred. No. 0.015;
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 QY 2 QVTDHLEHYLVEMQWTCCKP 23
 77 QITDMEHYIDVOISRNCCKP 98
 Db
 RESULT 5
 CST8 HUMAN STANDARD; PRT; 142 AA.
 ID CST8_HUMAN
 AC O60676;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
 related epididymal spermatogenic protein)
 GN CST8 OR CRGS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=95344753; PubMed=7619504;
 RA Cornwall G.A., Hann S.R.,
 RT "Transient appearance of CRGS protein during spermatogenesis and
 RT caput epididymal sperm maturation.";
 RL Mol. Reprod. Dev. 41:37-46(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=1638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Gardner C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lhvasaisho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McEay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehara H.K., Showkneen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

```

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Collinson A., Coville G.J., Deedman R., Dhant P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Kuy M.P., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kuehl P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lervasseth M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showheen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.B.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmington L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RA "The DNA sequence and comparative analysis of human chromosome 20."
RT Nature 414:865-871(2001).
RL -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- SIMILARITY: Belongs to the cystatin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1ab-sib.ch/announce/
CC or send an email to license@1ab-sib.ch).
CC -----
DR EMBL; AL096677; CAC03440.2; -.
DR HSSP; P01038; 1A90.
DR Genew; HGNC:15958; CSTL1.
DR InterPro; IPR000010; Cystatin.
DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 19
FT CHAIN 20 165
FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 111 121 BY SIMILARITY.
FT DISULFID 134 154
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 165 AA; 19312 MW; 9D66D85875DAEBA CRC64;
Query March 39.9%; Score 55; DB 1; Length 165;
Best Local Similarity 40.9%; Pred. No. 0.28;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Oy 1 RQVTDHLEYLHNLNVMQWTTQCK 22
Db 92 RQLTGVETIVTVKIGMTKCKR 113
RESULT 7
CYTC RAT STANDARD; PRT; 127 AA.
AC P14841.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Fragment).
GN CGT3.
OS Rattus norvegicus (Rat).

```

```

01 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
02 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
03 NCBI_TaxId=10116.
04
05 SEQUENCE FROM N.A.
06 [1]
07 STRAIN=Buffalo;
08 MEDLINE=90092122; PubMed=2689174;
09 Cole T., Dickson P.W., Esnard F., Averill F., Risbridger G.,
10 Gauchier F., Schneider G.;
11 "The cDNA structure and expression analysis of the genes for the
12 cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
13 in rat brain.";
14 Eur. J. Biochem. 186:35-42(1999).
15 [2]
16 SEQUENCE OF 8-127.
17 MEDLINE=90380276; PubMed=2400577;
18 Esnard F., Esnard A., Faucher D., Capony J.-P., Derancourt J.,
19 Brillard M., Gauchier F.;
20 "Rat cystatin C: the complete amino acid sequence reveals a site for
21 N-glycosylation.";
22 Biol. Chem. Hoppe-Seyler 371:161-166(1990).
23 [3]
24 SEQUENCE OF 8-49.
25 MEDLINE=88313020; PubMed=3044831;
26 Esnard A., Esnard F., Faucher D., Gauchier F.;
27 "Two rat homologues of human cystatin C.";
28 FEBS Lett. 236:475-478(1988).
29 [4]
30 SEQUENCE OF 8-20.
31 TISSUE=Sertoli cells;
32 MEDLINE=92225121; PubMed=1563513;
33 Esnard A., Esnard F., Guillou F., Gauchier F.;
34 "Production of the cysteine proteinase inhibitor cystatin C by rat
35 Sertoli cells.";
36 FEBS Lett. 300:131-135(1992).
37 -I- FUNCTION: As an inhibitor of cysteine proteases, this protein is
38 thought to serve an important physiological role as a local
39 regulator of this enzyme activity. Known to inhibit cathepsin B,
40 H, and L.
41 -----
42 -I- SIMILARITY: Belongs to the cystatin family.
43 -----
44 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
45 between the Swiss Institute of Bioinformatics and the EMBL outstation -
46 the European Bioinformatics Institute. There are no restrictions on its
47 use by non-profit institutions as long as its content is in no way
48 modified and this statement is not removed. Usage by and for commercial
49 entities requires a license agreement (See http://www.isb-sib.ch/announce/
50 or send an email to license@isb-sib.ch).
51 -----
52 DR EMBL; X16957; GAA34831.1; -.
53 DR PIR; S07085; S07085.
54 DR PIR; S10587; S10587.
55 DR HSSP; P01034; I636.
56 InterPro; IPR00010; Cystatin.
57 Pfam; PF00031; cystatin; 1.
58 SMART; SM00043; CY; 1.
59 PROSITE; PS00287; CYSTATIN; 1.
60 KW Thiol protease inhibitor; Signal.
61 FT NON TER 1 1
62 FT SIGNAL <1 7
63 FT CHAIN 8 127 CYSTATIN C.
64 FT ACT SITE 18 18 REACTIVE SITE.
65 FT SITE 62 66 SECONDARY AREA OF CONTACT.
66 FT DISULFID 80 90 BY SIMILARITY.
67 FT DISULFID 104 124 BY SIMILARITY.
68 FT CONFLICT 25 25 A -> E (IN REF. 2).
69 SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;
70
71 Query Match 38.4%; Score 53; DB 1; Length 127;
72 Best Local Similarity 41.7%; Pred. No. 0.43;
73 Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0

```

```

Db          : 61 KQVAGINYYLIDVEMGRITCKRSQ 84

RESULT 8
CYTC_BOVIN  STANDARD;  PRT;  148 AA.
ID_CYTC_BOVIN  P01035;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).
OS Cst3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NC NCB1_TaxID=9913;
[1]
SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.
RC TISSUE=Cerebrospinal fluid, AND Choroid plexus;
RX MEDLINE=98094199; PubMed=9434110;
RA Olsson S.-L., Ek B., Wilm M., Broberg S., Rask L., Bjork I.;
RT "Molecular cloning and N-terminal analysis of bovine cystatin C
RT identification of a full-length N-terminal region.";
RN Biochim. Biophys. Acta 1343:203-210(1997).
[2]
SEQUENCE OF 37-148.
RX MEDLINE=85231205; PubMed=3891407;
RA Hitado M., Terasawa S., Sakiyama F., Ninobe M., Fujii S.;
RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine
RT proteinase inhibitor.";
RN FEBS Lett. 186:41-45(1985).
[1] FUNCTION: This is a thiol proteinase inhibitor.
[1] MASS SPECTROMETRY: MW=13420; METHOD=MALDI.
[1] SIMILARITY: Belongs to the cystatin family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb.ch).
-----
CC EMBL; Y10811; CAJ71771.1; -.
CC HSSP; P01034; IG96.
DR InterPro: IPR000100; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
DR K1M Thiol proteinase inhibitor; Signal, Pyrrolidone carboxylic acid.
FT SIGNAL 1 30 PROBABLE.
FT CHAIN 1 148 CYSTATIN C.
FT MOD RES 31 31 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT ACT SITE 40 40 REACTIVE SITE.
FT SITE 84 88 SECONDARY AREA OF CONTACT.
FT DISULFID 102 112 BY SIMILARITY.
FT DISULFID 126 146 BY SIMILARITY.
SQ SEQUENCE 148 AA; 16265 MW; E8740FE37CFE9F0E CRC64;

Query Match 37.7%; Score 52; DB 1; Length 148;
Best Local Similarity 41.7%; Pred. No. 0.71;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

CY 1 KQVLDHELYHNEMQWTTCKRPE 24
Db 83 KQVVGSGMYFLDVELGRITCKRSQ 106

RESULT 9
CYTC_SAISC
ID_CYTC_SAISC  STANDARD;  PRT;  146 AA.

```

```

AC 019093;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN Cst3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C, Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC -----
DR EMBL; U52028; AAB64051.1; -.
DR HSSP; P01034; 1G96.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy.1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 146 CYSTATIN C.
FT ACT SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15946 MW; 08196353C0306AA3 CRC64;

Query March 36.2%; Score 50; DB 1; Length 146;
Best Local Similarity 45.5%; Pred. No. 1.4;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 RQVTDHLEVHLNVEVMTTCOK 22
   : : : : :
Db 80 KQIVAGVNYFLDVEVMTTCOK 101

RESULT 10
CYTC MOUSE STANDARD; PRT; 140 AA.
ID CYTC MOUSE
AC P21460;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Cystatin 3).
GN Cst3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=91054522; PubMed=2241983;
RX Solem W., Rawson C., Lindburg K., Barnes D.;
RA "Transforming growth factor beta regulates cystatin C in serum-free

```

```

RT mouse embryo (SEME) cells.";
RN Biochem. Biophys. Res. Commun. 172:945-951(1990).
RL [2]
RP SEQUENCE FROM N.A.
RX STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=95137392; PubMed=7835704;
RA Huh C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
RT "Structural organization, expression and chromosomal mapping of the
RT mouse cystatin-C-encoding gene (Cst3).";
RL Gene 152:221-226 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=ILIS, and ISS;
RX MEDLINE=21363810; PubMed=11471062;
RA Ehlinger M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikeja J.M.;
RT "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs.";
RL Mamm. Genome 12:657-663 (2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Datchenko L., Marzula K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Canninci P., Krangel C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hellton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallue D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ibt-sib.ch/announce/
CC or send an email to license@ibt-sib.ch).
CC -----
DR EMBL; M59470; AAA63298.1; -.
DR EMBL; U10098; AAB41056.1; -.
DR EMBL; AF483486; AAL90760.1; -.
DR EMBL; AF483487; AAL90761.1; -.
DR EMBL; BC002072; AAH02072.1; -.
DR PIR; A36163; A36163.
DR HSSP; P01034; 1G96.
DR MGD; MGI:102519; Cst3.
DR InterPro; IPR00010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 20
FT CHAIN 21 140 CYSTATIN C.
FT ACT SITE 31 31 REACTIVE SITE.
FT SITE 75 79 SECONDARY AREA OF CONTACT.
FT DISULFID 93 103 BY SIMILARITY.

```



```

FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 110 BY SIMILARITY.
FT DISULFID 124 144 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 145 AA; 16454 MW; B2BCA4F6857CB0F CRC64;

Query Match 35.5%; Score 49; DB 1; Length 145;
Best Local Similarity 47.6%; Pred. No. 2;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 QVTDHLEHLNVEWQMTTCOK 22
Db 81 QIVAGLKMYLEVEIGRTTCKK 101

RESULT 12
CYT_BITAR STANDARD; PRT; 111 AA.
ID_CYT_BITAR
AC P08935;
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin.
OS Bittis arletans (African puff adder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Bittis.
OX NCBI_TaxID=8692;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=88076861; PubMed=3500714;
RA Ritonja A., Evans H.J., Machleidt W., Barrett A.J.;
RT "Amino acid sequence of a cystatin from venom of the African puff
RT adder (Bittis arletans).";
RL Biochem. J. 246:799-802(1987).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the cystatin family.
DR HSSP; A28793; A28793.
DR HSSP; P01038; ICBW.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KM Thiol protease inhibitor.
FT ACT_SITE 3 3 REACTIVE SITE.
FT SITE 47 51 SECONDARY AREA OF CONTACT.
FT DISULFID 65 81 PROBABLE.
FT VARIANT 57 57 M -> T (IN EQUAL AMOUNT).
SQ SEQUENCE 111 AA; 12678 MW; 1A3B6B246AC10C CRC64;

Query Match 34.8%; Score 48; DB 1; Length 111;
Best Local Similarity 42.9%; Pred. No. 2.1;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHLNVEWQMTTCOK 22
Db 47 QVVGSKRYLMMLKTTCKK 67

RESULT 13
CYT_CHICK STANDARD; PRT; 139 AA.
ID_CYT_CHICK
AC P01036;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin precursor (Egg-white cystatin).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```

```

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90008673; PubMed=2793849;
RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;
RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
RT and tissue distribution.";
RL J. Biol. Chem. 264:17164-17169(1989).
RN [2]
RP SEQUENCE OF 24-139
RX MEDLINE=84178305; PubMed=6712597;
RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
RT "Cystatin. Amino acid sequence and possible secondary structure.";
RL Biochem. J. 217:813-817(1984).
RN [3]
RP SEQUENCE OF 24-139
RX MEDLINE=84110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
RA Machleidt W.;
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
RT of cystatin from chicken egg white.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [4]
RP CHARACTERIZATION OF PROTEIN.
RX MEDLINE=83256421; PubMed=6409085;
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
RA Sauter D.C., Barrett A.J.;
RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
RT purification from egg white, characterization, and detection in
RT chicken serum.";
RL Biochem. J. 211:129-138(1983).
RN [5]
RP DISULFIDE BONDS.
RX Grubb A., Loeberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
RT cystatin.";
RL FEBS Lett. 170:370-374(1984).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=89252033; PubMed=2721673;
RA Laber B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,
RA Bode W.;
RT "The cysteine proteinase inhibitor chicken cystatin is a
RT phosphoprotein.";
RL FEBS Lett. 248:162-168(1989).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=89052676; PubMed=3191914;
RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshkov A.,
RA Brzin J., Kos J., Turk V.;
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
RT its possible mode of interaction with cysteine proteinases.";
RL EMBO J. 7:2593-2599(1988).
RN [8]
RP STRUCTURE BY NMR.
RX MEDLINE=94087719; PubMed=8263912;
RA Dieckmann T., Mitschang U., Hofmann M., Kos J., Turk V.,
RA Auerwald E.A., Jeanicke R., Oeschkinat H.;
RT "The structures of native phosphorylated chicken cystatin and of a
RT recombinant unphosphorylated variant in solution.";
RL J. Mol. Biol. 234:1048-1059(1993).
CC -1- FUNCTION: This protein binds tightly to and inhibits a variety of
CC thiol proteases including ficin, papain, and cathepsins B, C, H,
CC and L. Although isolated from egg white, it is also present in
CC serum.
CC -1- SIMILARITY: Belongs to the cystatin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: J03077, AAA48744.1; -.

DR PIR: A34456; UDCH.

DR PDB: 1CEW; 31-JAN-94.

DR PDB: 1A67; 27-MAY-98.

DR PDB: 1A90; 17-JUN-98.

DR InterPro: IPR000010; Cystatin.

DR Pfam: PF00031; cystatin.1.

DR SMART: SM00043; CY.1.

DR PROSITE: PS00287; CYSTATIN.1.

KM Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.

FT SIGNAL 1 23

FT CHAIN 24 139

FT ACT_SITE 32 32

FT SITE 76 80

FT DISULFID 94 104

FT DISULFID 118 138

FT MOD_RES 103 103

FT STRAND 35 36

FT TURN 39 40

FT TURN 42 51

FT TURN 52 52

FT HELIX 53 56

FT TURN 57 58

FT STRAND 63 77

FT TURN 81 95

FT TURN 96 97

FT TURN 99 100

FT HELIX 101 108

FT STRAND 115 125

FT TURN 126 129

FT STRAND 130 139

SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;

Query March 34.8%; Score 48; DB 1; Length 139;

Best Local Similarity 45.5%; Pred. No. 2.7;

Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 RQVTDHLEHNLVEMQWTTCK 22

Db 75 RQLVSGIKYILQVEIGRTTCK 96

RESULT 14

CYTC HUMAN

AC CYTC HUMAN STANDARD; PRT; 146 AA.

ID P01034;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)

DE (Post-gamma-globulin).

GN CSTR3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;

RT "Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C.;"

RL FEBS Lett. 216:229-233(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Leukocyte;

RA Abrahamson M., Olafsson I., Paledottir A., Ulvbaeck M., Lundwall A.,

RA Jensen O., Grubb A.;

RT "Structure and expression of the human cystatin C gene.;"

RL Biochem. J. 268:287-294(1990).

RN [3]

RP SEQUENCE FROM N.A. (HCMA VARIANT).

RC TISSUE=Brain;

RA MEDLINE=89235594; PubMed=2541223;

RA Levy E., Lopez-Otin C., Ghiso J., Gellner D., Frangione B.;

RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a mutation in the cystatin C gene, an inhibitor of cysteine proteases.;"

RT J. Exp. Med. 169:1771-1778(1989).

RL [4]

RP SEQUENCE FROM N.A.

RA MEDLINE=89350949; PubMed=2764935;

RA Satoh E., Sabatini L.M., Eddy R.L., Snows T.B., Azen E.A.,

RA Isemura S., Sanada K.;

RT "The human cystatin C gene (CSTR3) is a member of the cystatin gene family which is localized on chromosome 20.;"

RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).

RN [5]

RP SEQUENCE FROM N.A.

RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;

RT "Acquisition of complex patterns of differential expression in epithelial cell populations during the evolution of type 2 cystatin genes.;"

RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

RL [6]

RP SEQUENCE FROM N.A.

RA MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Cleghorn S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights K., Laydon G.K., Lawlor S.,

RA Lehaesliho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie I.J., McElay K., McMurtry A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,

RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,

RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,

RA Whitehead S.L., Whitaker P., Willey D.L., Williams S., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.;"

RL Nature 414:865-871(2001).

RN [7]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA MEDLINE=22388257; PubMed=12477932;

RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Butow K.H., Schaefer C.F., Bhac N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,

RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,

RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,

RA Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,

RA Villalón D.K., Muzny K.C., Dalm S., Sodergren B.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RN SEQUENCE OF 27-146.
 RX MEDLINE=8222268; PubMed=6283552;
 RA Grubb A., Loefberg H.;
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and
 presence in the adenohypophysis";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).
 RN [9]
 RN SEQUENCE OF 27-73.
 RX MEDLINE=84110059; PubMed=6662498;
 RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
 RA Machleidt W.;
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 of cystatin from chicken egg white";
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
 RN [10]
 RN SEQUENCE OF 27-76.
 RX MEDLINE=84128015; PubMed=6365094;
 RA Brzin J., Popovic T., Turk V.;
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases";
 RL Biochem. Biophys. Res. Commun. 118:103-109(1984).
 RN [11]
 RN DISULFIDE BONDS.
 RA Grubb A., Loefberg H., Barrett A.J.;
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 cystatin";
 RL FEBS Lett. 170:370-374(1984).
 RN [12]
 RN X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.
 RX MEDLINE=21173909; PubMed=11276250;
 RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,
 RA Abrahamson M., Jaskolski M.;
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through
 three-dimensional domain swapping";
 RL Nat. Struct. Biol. 8:316-320(2001).
 RN [13]
 RN VARIANT GLN-94.
 RX MEDLINE=92316504; PubMed=1352269;
 RA Abrahamson M., Jonedottr S., Olafsson I., Jansson O., Grubb A.;
 RT "Hereditary cystatin C amyloid angiopathy: identification of the
 disease-causing mutation and specific diagnosis by polymerase chain
 reaction based analysis";
 RL Hum. Genet. 89:377-380(1992).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 thought to serve an important physiological role as a local
 regulator of this enzyme activity.
 CC -1- SUBUNIT: Homodimer.
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,
 vas deferens, brain, thymus, and ovary and the lowest in the
 submandibular gland.
 CC -1- DISEASE: Defects in CST3 are a cause of hereditary cerebral
 hemorrhage with amyloidosis (HCHWA) [MM:105150]; also known as
 cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis
 Icelandic type. HCHWA is characterized by a thickening of the
 cerebral arteries walls with deposition of material with the
 characteristics of amyloid.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X05607; CAA29096.1; -;
 DR EMBL; X52255; CAA36497.1; -;
 DR EMBL; M27891; AAA52164.1; -;
 DR EMBL; M27889; AAA52164.1; JOINED.
 DR EMBL; M27890; AAA52164.1; JOINED.
 DR EMBL; X61681; CAA43856.2; -;
 DR EMBL; X61682; CAA43856.2; JOINED.
 DR EMBL; AF319564; AAK11570.1; -;
 DR EMBL; AL121894; CAC05424.1; -;
 DR EMBL; BC013083; AAH13083.1; -;
 DR PIR; S10216; UDHU.
 DR PDB; 1G96; 06-APR-01.
 DR Genem; HGNC:2475; CST3.
 DR MIM; 604312; -;
 DR MIM; 105150; -;
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;
 KW Polymorphism; 3d-structure.
 FT SIGNAL 1 26
 FT CHAIN 27 146
 FT ACT SITE 37 37
 FT SITE 81 85
 FT DISULFID 99 109
 FT DISULFID 123 143
 FT
 Query Match 34.8%; Score 48; DB 1; Length 146;
 Best Local Similarity 37.5%; Pred. No. 2.8;
 Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 Oy 1 RQVTDHEFLHNVEMQWTTCKRPE 24
 Db 80 KQIVAGVYFLDVELGRITCTKTQ 103
 ID CYTC_MACMU STANDARD; PRT; 146 AA.
 AC O19092;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor.
 GN CST3.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=97054523; PubMed=8898820;
 RA Wei L.H., Walker L.C., Levy E.;
 RT "Cystatin C, Icelandic-like mutation in an animal model of
 cerebrovascular beta-amyloidosis";
 RL Stroke 27:2080-2085(1996).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 thought to serve an important physiological role as a local
 regulator of this enzyme activity.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U51912; AAB64050.1; -;

DR HSSP: P01034; 1G96.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 KM Thiol protease inhibitor; Amyloid; Signal.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 146 CYSTATIN C.
 FT ACT_SITE 37 37 REACTIVE SITE.
 FT SITE 81 85 SECONDARY AREA OF CONTACT.
 FT DISULFID 99 109 BY SIMILARITY.
 FT DISULFID 123 143 BY SIMILARITY.
 SQ SEQUENCE 146 AA; 15857 MW; F083BB774A29DF26 CRC64;

Query Match 34.8%; Score 48; DB 1; Length 146;
 Best Local Similarity 37.5%; Pred. No. 2.8;
 Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 ROYTDHLEHYHNVEMQWTCOKPE 24
 Db 80 KQIVAGVNFPLDVELGRITCKTQ 103

Search completed: March 18, 2004, 14:16:12
 Job time : 3.85714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 14.381 Seconds
(without alignments)
526.560 Million cell updates/sec

Title: US-09-941-314-11
Perfect score: 138
Sequence: 1 RQVTDHLEHYHNVEMQWTCCKPE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	138	4	Q8WXU6
2	84	60.9	139	11	Q8K5A3
3	60	43.5	141	11	Q9DAP1
4	60	43.5	141	11	Q80ZM5
5	57	41.3	236	16	Q8DS42
6	57	41.3	236	16	Q8DS42
7	53	38.4	1779	9	Q94M04
8	52	37.7	540	5	Q8T9G8
9	52	37.7	966	5	Q8T9K4
10	52	37.7	966	5	Q9VLM8
11	52	37.7	966	5	Q9VLM8
12	51.5	37.3	678	2	Q83XK8
13	51.5	37.3	2165	5	Q8T1F5
14	50	36.2	232	12	Q9YVNA
15	50	36.2	588	16	Q97N97
16	50	36.2	588	16	Q8DN24

17	49.5	35.9	256	9	Q7Y4K3	Q7Y4K3 streptococ
18	49	35.5	140	11	Q9BPX9	Q9BPX9 mus musculu
19	49	35.5	148	11	Q9JW84	Q9JW84 mus musculu
20	49	35.5	167	4	Q7Z4U8	Q7Z4U8 homo sapien
21	48.5	35.1	3546	2	Q9F8J0	Q9F8J0 micromonos
22	48	34.8	278	16	Q8LVI4	Q8LVI4 bacillus an
23	48	34.8	278	16	Q8LVB4	Q8LVB4 bacillus ce
24	48	34.8	1078	10	Q94LJ7	Q94LJ7 oryza sativ
25	48	34.8	1078	10	Q7XH40	Q7XH40 oryza sativ
26	47.5	34.4	302	4	Q9H740	Q9H740 homo sapien
27	47.5	34.4	421	4	Q9NXX0	Q9NXX0 homo sapien
28	47.5	34.4	544	6	Q28548	Q28548 ovis aries
29	47.5	34.4	610	4	Q9UJZ2	Q9UJZ2 homo sapien
30	47.5	34.4	631	4	Q9NXX3	Q9NXX3 homo sapien
31	47.5	34.4	693	11	Q8K145	Q8K145 mus musculu
32	47.5	34.4	737	4	Q8ND82	Q8ND82 homo sapien
33	47.5	34.4	787	11	Q8B182	Q8B182 mus musculu
34	47.5	34.4	935	5	Q18671	Q18671 caenorhabdi
35	47.5	34.4	995	4	Q9HC18	Q9HC18 homo sapien
36	47.5	34.4	1308	17	Q9VXG5	Q9VXG5 pyrococcus
37	47	34.1	273	16	Q8ZCK5	Q8ZCK5 streptomyce
38	47	34.1	392	5	Q9VY83	Q9VY83 drosophila
39	47	34.1	1139	5	Q9VW15	Q9VW15 drosophila
40	46.5	33.7	199	10	Q9FPG2	Q9FPG2 arabidopsis
41	46.5	33.7	367	10	Q9CGR6	Q9CGR6 arabidopsis
42	46.5	33.7	371	10	Q9CSP7	Q9CSP7 arabidopsis
43	46.5	33.7	678	2	Q83XU6	Q83XU6 legionella
44	46.5	33.3	79	15	Q07406	Q07406 simian t-ly
45	46	33.3	167	11	Q9QWL5	Q9QWL5 mus musculu

ALIGNMENTS

RESULT 1
ID Q8WXU6 PRELIMINARY; PRT; 138 AA.

AC Q8WXU6; 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE SC13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]

RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French P.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF35480; AAL7191.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; CY; 1.
SQ SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;

Query Match 100.0%; Score 138; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQVTDHLEHYHNVEMQWTCCKPE 24
Db 75 RQVTDHLEHYHNVEMQWTCCKPE 98

RESULT 2
ID Q8K5A3 PRELIMINARY; PRT; 139 AA.
AC Q8K5A3;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
 DE Cystatin 11.
 GN CST11.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Hamil K.G., Hall S.H.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF501290; AAM21709.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 139 AA; 16686 MW; E1E36DB786B4D08C CRC64;

Query Match 60.9%; Score 84; DB 11; Length 139;
 Best Local Similarity 58.3%; Pred. No. 2.2e-05;
 Matches 14; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 QVTDHLEHYLVNEMQWTTCKP 24
 DB 75 KQMTNMEFHITVEMORTTCKTE 98

RESULT 3

Q9DAP1 PRELIMINARY; PRT; 141 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
 DE 170006C19Rik protein.
 GN 170006C19Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fleischer C., Fujita M., Gariboldi M.F.,
 RA Guetincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombereis P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Winking L.,
 RA Wysshaw-Bois A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
 RA Hayaishiaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK005665; BAB24175.1; -
 DR HSSP; P01038; ICEW.
 DR MGD; MGI:1916544; 170006C19Rik.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8A1AC378C CRC64;

Query Match 43.5%; Score 60; DB 11; Length 141;
 Best Local Similarity 52.4%; Pred. No. 0.14;
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

RESULT 4

Q80ZNS PRELIMINARY; PRT; 141 AA.

DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 25, last annotation update)
 DE RIKEN cDNA 170006C19 gene.
 GN 170006C19.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testicle;
 RA Straubeberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC048681; AAH48681.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR ProDom; PD001231; Cystatin_C/M; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8A884951F CRC64;

Query Match 43.5%; Score 60; DB 11; Length 141;
 Best Local Similarity 52.4%; Pred. No. 0.14;
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 QVTDHLEHYLVNEMQWTTCK 22
 DB 76 QITDSLEYLVNARTMCKK 96

RESULT 5

Q8DS42 PRELIMINARY; PRT; 236 AA.

DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
 DE Putative ABC transporter, ATP-binding protein.
 GN ABC OR SMU.1994.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Ikin S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferrerelli J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE015022; AAN59598.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.


```
DR GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
KM ATP-binding; Complete proteome.
SQ SEQUENCE 236 AA; 26812 MW; 1BAB8DD742DC76DB CRC64;

Query Match 41.3%; Score 57; DB 16; Length 236;
Best Local Similarity 37.0%; Pred. No. 0.7;
Matches 10; Conservative 6; Mismatches 7; Indels 4; Gaps 1;

Qy 1 RQVTHLEHYLNVEMQ-----WTCQKP 23
Db 108 RRLTKHDDHDKVLSVGMWENCQKP 134

RESULT 6
Q94M04 PRELIMINARY; PRT; 723 AA.
ID 094M04
AC 094M04;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Major core structural protein pI.
GN 1.
OS Bacteriophage phi-12.
OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
OC NCBI_TaxID=161736;
RN (1)
RP SEQUENCE FROM N.A.
RA Gottlieb P.J., Potgieter C., Wei H., Toporovsky I.;
RT "Characterization of Bacteriophage phi12."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF08636; AL001109.1; -.
DR InterPro: IPR006162; Pphanth_S.
DR PROSITE: PS00012; PHOSPHOPANTHINE; 1.
SQ SEQUENCE 723 AA; 78950 MW; F5DEF7FF3A50B7C CRC64;

Query Match 41.3%; Score 57; DB 9; Length 723;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 6 HLEHYLNVEMQWTCQKP 23
Db 463 HTEYRFDVLCWATCYP 480

RESULT 7
Q18150 PRELIMINARY; PRT; 1779 AA.
ID 018150
AC 018150; O45627;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE T28B8.4 protein.
GN T28B8.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
```

```
RA Craxton M., Dear S., Du Z., Durbin R., Pavello A., Fulton L.,
RA Gardner A., Green P., Hawkes T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten T., Laister N., Latreille P.,
RA Lightning U., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Showkeen R.,
RA Samalson N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sprout J., Woldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38 (1994).
RN (3)
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81133; CAB03445.1; -.
DR EMBL; AL021066; CAB03445.1; JOINED.
DR EMBL; AL021066; CAI5925.1; -.
DR EMBL; Z81133; CAI5925.1; JOINED.
DR PIR; T23130; T23130.
DR WormPep; T28B8.4; CB16519.
DR InterPro: IPR008938; ANM.
SQ SEQUENCE 1779 AA; 205511 MW; 70A4489148B96F38 CRC64;

Query Match 38.4%; Score 53; DB 5; Length 1779;
Best Local Similarity 37.0%; Pred. No. 26;
Matches 10; Conservative 5; Mismatches 4; Indels 8; Gaps 1;

Qy 1 RQVTHLEHYLNVEMQWTCQKP 19
Db 749 RRLTKHDDHDKVLSVGMWENCQKP 775

RESULT 8
Q8T9G8 PRELIMINARY; PRT; 540 AA.
ID 08T9G8
AC 08T9G8;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE SD01519P.
GN AATS-ALA OR CG13391.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069773; AL039918.1; -.
DR FLYbase; FBgn0027094; Aats-ala.
DR GO; GO:0004813; F:alanine-tRNA ligase activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
DR InterPro: IPR003156; DHHA1.
DR InterPro: IPR002318; CRNA-synt_2c.
DR InterPro: IPR006193; CRNA-synt_Ala.
DR Pfam; PF02272; DHHA1; 1.
DR Pfam; PF01411; CRNA-synt_2c; 1.
DR PROSITE: PS50860; AA TRNA_LIGASE II ALA; 1.
SQ SEQUENCE 540 AA; 59580 MW; B670DCF9C38684B CRC64;
```

```
Query Match 37.7%; Score 52; DB 5; Length 540;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

QY 2 QVTDHLEHYHLNVEMQWTT 19
 Db 157 KVGDELEHLIDVERRWLT 174

RESULT 9

ID Q879K4 PRELIMINARY; PRT; 966 AA.
 AC Q879K4;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE GN03058P.
 GN AATS-ALA OR CG13391.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Munoz J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY069255; AL339400.1; -
 DR FLYBase; FBgn0027094; Aats-ala.
 DR GO; GO:0004813; F:alanine-cRNA ligase activity; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0006419; F:nucleic acid binding; IEA.
 DR GO; GO:0006419; P:alanyl-cRNA aminoacylation; IEA.
 DR InterPro; IPR003156; DHHA1.
 DR InterPro; IPR002318; cRNA-synt_2c.
 DR InterPro; IPR006193; cRNA-synt_Ala.
 DR Pfam; PF02272; DHHA1; 1.
 DR Pfam; PF01411; cRNA-synt_2c; 1.
 DR PRINTS; PR00980; TRNASYNTHALA.
 DR TIGRFAMs; TIGR00344; alas; 1.
 DR PROSITE; PSS0860; AA TRNA_LIGASE_II_ALA; 1.
 DR SEQUENCE 966 AA; 107697 MW; 9F2079A119F463B4 CRC64;
 SQ
 QY 2 QVTDHLEHYHLNVEMQWTT 19
 Db 583 KVGDELEHLIDVERRWLT 600
 Query Match 37.7%; Score 52; DB 5; Length 966;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bakendale U., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Buritis K.C., Buzam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dukov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ilegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkovic R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003621; AATF52657.1; -
 DR FLYBase; FBgn0027094; Aats-ala.
 DR GO; GO:0004813; F:alanine-cRNA ligase activity; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0006419; F:nucleic acid binding; IEA.
 DR GO; GO:0006419; P:alanyl-cRNA aminoacylation; IEA.
 DR InterPro; IPR003156; DHHA1.
 DR InterPro; IPR002318; cRNA-synt_2c.
 DR InterPro; IPR006193; cRNA-synt_Ala.
 DR Pfam; PF02272; DHHA1; 1.
 DR Pfam; PF01411; cRNA-synt_2c; 1.
 DR PRINTS; PR00980; TRNASYNTHALA.
 DR TIGRFAMs; TIGR00344; alas; 1.
 DR PROSITE; PSS0860; AA TRNA_LIGASE_II_ALA; 1.
 DR SEQUENCE 966 AA; 107741 MW; B1B63A7A8FB87B6E CRC64;
 SQ
 QY 2 QVTDHLEHYHLNVEMQWTT 19
 Db 583 KVGDELEHLIDVERRWLT 600
 Query Match 37.7%; Score 52; DB 5; Length 966;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

RESULT 11

ID Q9U6B4 PRELIMINARY; PRT; 966 AA.
 AC Q9U6B4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Alanyl-cRNA synthetase.
 GN AATS-ALA OR CG13391.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

```

OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RA Chihade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.;
RT "Detection of an Intermediate Stage of Mitochondria Genesis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188718; AAF05593.1; -.
DR FlyBase; FBgn0027094; Aats-ala.
DR GO; GO:0004813; P:alanine-tRNA ligase activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR002318; tRNA-synt-2c.
DR InterPro; IPR006193; tRNA-synt-Ala.
DR Pfam; PF02272; DHHA1; 1.
DR Pfam; PF01411; tRNA-synt-2c; 1.
DR PRINTS; PR00980; TRNASYNTHALA.
DR TIGRPFAM; TIGR00344; alas; 1.
DR PROSITE; PSS0860; AA_TRNA_LIGASE_II_ALA; 1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 966 AA; 107877 MW; 82034A189F0C81E5 CRC64;

Query Match 37.7%; Score 52; DB 5; Length 966;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QVTDHLEYHLNVEMQWTT 19
Db 583 KVGDELEHLHIDVERHWLT 600

RESULT 12
Q83XK8 PRELIMINARY; PRT; 678 AA.
ID Q83XK8;
AC Q83XK8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lezy protein.
GN LSSY.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Philadelphia;
RA Heuner K.;
RT "Description of a putative type I secretion system of Legionella
RT pneumophila.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ560770; CAD90962.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008934; AcPase_VanPase.
DR InterPro; IPR000252; Deda.
DR InterPro; IPR000326; PA_PTPase.
DR Pfam; PF00597; Deda; 1.
DR Pfam; PF01569; PAP2; 1.
SQ SEQUENCE 678 AA; 76895 MW; AC817B2813F03DAF CRC64;

Query Match 37.3%; Score 51.5; DB 2; Length 678;
Best Local Similarity 44.0%; Pred. No. 16;
Matches 11; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

Qy 2 QVTDHLEYHLNVEMQWTTCKP 23
Db 480 QVHDHMPYHKEFTLTETKQMDQCP 504

RESULT 13
Q8T1F5 PRELIMINARY; PRT; 2165 AA.
ID Q8T1F5

```

```

AC Q8T1F5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Dictyostellium discoideum (Slime mold). protein tyrosine
DE kinase.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostellium.
OX NCBI_TaxID=44689;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RL MBLINB-22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szefranek K., Pachet J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostellium discoideum.";
RN Nature 418:79-85 (2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116959; AAM08443.2; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001683; PK_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD0000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PSS0003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PSS0195; PK; 1.
KW Kinase.
SQ SEQUENCE 2165 AA; 241910 MW; AFFF8B5B1C32202F CRC64;

Query Match 37.3%; Score 51.5; DB 5; Length 2165;
Best Local Similarity 37.5%; Pred. No. 55;
Matches 9; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

Qy 2 QVTDHLEYHLNVEMQWTTCKP 22
Db 658 QIDDHLEYKQVLSNSNEWFCRR 681

RESULT 14
Q9YVN4 PRELIMINARY; PRT; 232 AA.
ID Q9YVN4;
AC Q9YVN4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF MSV208 putative uracil DNA glycosylase UNG (vaccinia DAR), similar
DE to Equine herpesvirus GB/U20824.
GN MSV208.
OS Melanoplus sanguinipes entomopoxvirus (MSFPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_TaxID=83191;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson;

```

RX MEDLINE=99102612; PubMed=9847359;
RA Alonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;
RT "The genome of Melanoplus sanguinipes entomopoxvirus";
RN J. Virol. 73:533-552(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson;
RA Alonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF063866; AAC9753.1; -.
DR PIR; T28369.
DR GO; GO:0004844; F:uracil DNA N-glycosylase activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR005122; UDNA_glycylaseF.
DR InterPro; IPR003249; U_glycylase_notp.
DR Pfam; PF03167; UDG; 1.
DR PRODOM; PD001589; U_glycylase_notp; 1.
SQ SEQUENCE 232 AA; 27297 MW; 85A9879E1C4F09A0 CRC64;

Query Match 36.2%; Score 50; DB 12; Length 232;
Best Local Similarity 31.6%; Pred. No. 8.8;
Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Oy 2 QVTDHLEVHNVEMQWTTTC 20
Db 212 KANEYKIKHKNIEIDWISC 230

RESULT 15
O97N97
ID Q97N97 PRELIMINARY; PRT; 588 AA.
AC Q97N97;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE L-fucose isomerase.
GN SP2158.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RA MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Khouli H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
DR EMBL; AE007504; AAK76212.1; -.
DR PIR; C95252; C95252.
DR TIGR; SP2158; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0008736; F:L-fucose isomerase activity; IEA.
DR GO; GO:0006004; P:fucose metabolism; IEA.
DR InterPro; IPR004216; Fuc_isomeraseC.
DR InterPro; IPR009015; Fuc_isomeraseN.
DR Pfam; PF02952; fucose_180_C; 1.
SQ SEQUENCE 588 AA; 65894 MW; 06E678EDF801786A CRC64;

Query Match 36.2%; Score 50; DB 16; Length 588;
Best Local Similarity 37.9%; Pred. No. 24;
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Oy 1 RQVTDH-----LEVHNVEMQWTTTCOKP 23
Db 301 RQVTDHFPNGDFMETFLNTQFDWNGIRKP 329

Search completed: March 18, 2004, 14:21:30
Job time : 16.381 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 29.9226 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-12
Perfect score: 189
Sequence: 1 QVTDHLEHYLNVEMQWTCQKERTNCVPERE 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	33	AAU79862	AAU79862 Human cys
2	189	100.0	52	AAU79864	AAU79864 Human cys
3	189	100.0	59	AAU79866	AAU79866 Human cys
4	189	100.0	80	AAU79865	AAU79865 Human cys
5	189	100.0	115	AAU79853	AAU79853 Human cys
6	189	100.0	117	AAU79854	AAU79854 Human cys
7	189	100.0	137	AAU79852	AAU79852 Human cys
8	149	78.8	46	AAU79860	AAU79860 Human cys
9	149	78.8	49	AAU79863	AAU79863 Human cys
10	133	70.4	24	AAU79861	AAU79861 Human cys
11	133	70.4	48	AAU79867	AAU79867 Human cys
12	95	50.3	142	ADDA6708	ADDA6708 Rat Prote
13	95	50.3	142	ADDA6704	ADDA6704 Rat Prote
14	89	47.1	142	AAE02404	AAE02404 Murine cy
15	89	47.1	142	AAE04433	AAE04433 Mouse cys
16	89	47.1	143	ADA14374	ADA14374 Mouse spe
17	72.5	38.4	203	AAU83156	AAU83156 Novel sec
18	70	37.0	92	AAW78259	AAW78259 Fragment
19	70	37.0	133	AAW78260	AAW78260 Fragment
20	70	37.0	142	AAW78258	AAW78258 Fragment
21	70	37.0	142	AAE02405	AAE02405 Fragment
22	70	37.0	142	AAE04434	AAE04434 Human cys
23	70	37.0	142	ADA57231	ADA57231 Human sec
24	70	37.0	142	ADA41112	ADA41112 Human sec
25	70	37.0	142	ADC74335	ADC74335 Human sec

26	70	37.0	142	7	ADD37980	Add37980 Human sec
27	70	37.0	142	7	ADDA6706	ADDA6706 Human Pro
28	70	37.0	142	7	ADDA6710	ADDA6710 Human Pro
29	68.5	36.2	138	4	AAE65889	AAE65889 Amino aci
30	67.5	35.7	37	4	ABBA4210	ABBA4210 Peptide #
31	67.5	35.7	37	4	AAAM35913	AAAM35913 Peptide #
32	67.5	35.7	37	4	ABBA25688	ABBA25688 Protein #
33	67.5	35.7	37	4	AAW75805	AAW75805 Human bon
34	67.5	35.7	37	4	AAAM62992	AAAM62992 Human bra
35	67.5	35.7	37	4	ABG57540	ABG57540 Human liv
36	67.5	35.7	37	5	ABG45267	ABG45267 Human pep
37	67	35.4	141	4	AAAY96576	AAAY96576 Murine cy
38	67	35.4	141	4	AAE02403	AAE02403 Murine cy
39	67	35.4	141	4	AAE04432	AAE04432 Mouse tes
40	65	34.4	113	6	ADA57563	ADA57563 Human sec
41	65	34.4	113	6	ADA41457	ADA41457 Human sec
42	65	34.4	113	7	ADC74577	ADC74577 Human sec
43	65	34.4	113	7	ADD38088	ADD38088 Human sec
44	65	34.4	114	2	AAW78153	AAW78153 Human sec
45	65	34.4	127	7	ADB61282	ADB61282 Rat Prote

ALIGNMENTS

RESULT 1
AAU79862 standard; peptide; 33 AA.
XX
AC AAU79862;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcy8) antigenic fragment #10.
XX
KW Cystatin-8; Zcy8; cancer; procoagulant protein; thrombosis;
KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic peptide.
XX
OS Homo sapiens.
XX
PN W0200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026668.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-363044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 97; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcy8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis; modulating seminal
CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcy8 is useful as research reagent for characterizing sites
CC of interaction between Zcy8 and its receptor. Zcy8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcy8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (1) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 33 AA;

Query Match 100.0%; Score 189; DB 5; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHLNVMQWTTCKPPTNCVPOERE 33
 Db 1 QVTDHLEHLNVMQWTTCKPPTNCVPOERE 33

RESULT 2

AAU79864
 ID AAU79864 standard; peptide; 52 AA.

AC AAU79864;
 XX
 DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #12.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

PA (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis; modulating seminal
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(1) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (1) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 52 AA;

Query Match 100.0%; Score 189; DB 5; Length 52;
 Best Local Similarity 100.0%; Pred. No. 1.9e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHLNVMQWTTCKPPTNCVPOERE 33
 Db 20 QVTDHLEHLNVMQWTTCKPPTNCVPOERE 52

RESULT 3

AAU79866
 ID AAU79866 standard; peptide; 59 AA.

AC AAU79866;

XX
 DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #14.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.

XX Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

PA (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 99; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis; modulating seminal
 CC fluid viscosity; enhancing viability of cryopreserved sperm; sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(1) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (1) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (1) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX Sequence 59 AA;

Query Match 100.0%; Score 189; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 2.2e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33
 |||||
 Db 2 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 34

RESULT 4
 AAU79865
 ID AAU79865 standard; peptide; 80 AA.

AC AAU79865;
 XX
 DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #13.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.

OS Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 98; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

XX Sequence 80 AA;

Query Match 100.0%; Score 189; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 3.1e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33
 |||||
 Db 23 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 55

RESULT 5
 AAU79853
 ID AAU79853 standard; protein; 115 AA.

AC AAU79853;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #1.

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KW spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

XX WO200220567-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US026868.

XX 01-SEP-2000; 2000US-0230230P.

XX (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

XX WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

XX Claim 2; Page 94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)

CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.

CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.

CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)

XX Sequence 115 AA;

SO

Query Match 100.0%; Score 189; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 4.6e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 33
 DB 53 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 85

RESULT 6
 AAU79854
 ID AAU79854 standard; protein; 117 AA.
 AC AAU79854;
 XX
 XX 15-JUL-2002 (first entry)
 DE Human cystatin-8 (Zcys8) antigenic fragment #2.
 XX

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic fragment.
 XX
 OS Homo sapiens.
 XX WO200220567-A2.
 PN 14-MAR-2002.
 PD 29-AUG-2001; 2001WO-US026868.
 PF 01-SEP-2000; 2000US-0230230P.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Holloway JL, Gao Z, Bishop PD;
 PI WPI; 2002-383044/41.
 DR

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 94-95; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 CC
 XX
 SQ Sequence 117 AA;

Query Match 100.0%; Score 189; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 4.7e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 33
 DB 53 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 85

DB 55 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 87

RESULT 7
 AAU79852
 ID AAU79852 standard; protein; 137 AA.
 AC AAU79852;
 XX
 XX 15-JUL-2002 (first entry)
 DE Human cystatin-8 (Zcys8).
 XX

XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation.
 XX
 OS Homo sapiens.
 XX WO200220567-A2.
 PN 14-MAR-2002.
 PD 29-AUG-2001; 2001WO-US026868.
 PF 01-SEP-2000; 2000US-0230230P.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Holloway JL, Gao Z, Bishop PD;
 PI WPI; 2002-383044/41.
 DR N-PSDB; ABK49522.
 DR

XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 XX
 PS Claim 2; Page 93-94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This is the amino acid sequence of human cystatin-8 (Zcys8)
 CC
 XX
 SQ Sequence 137 AA;

Query Match 100.0%; Score 189; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 5.6e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 33
 DB 75 QVTDHLEHYHNVEMQWTTCKPFTTNCVPOERE 107


```

RESULT 8
AAU79860
ID AAU79860 standard; peptide: 46 AA.
XX
AC AAU79860;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8) antigenic fragment #8.
XX
KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
XX spermatoogenesis; seminal fluid viscosity; cryopreserved sperm;
XX sperm motility; fertilisation; antigenic peptide.
XX
OS Homo sapiens.
XX
PN WO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatoogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 97; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatoogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 46 AA;
XX
Query Match 78.8%; Score 149; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. NO. 5.2e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 QVTDHLEFYLNVEMQWTTCKPRTTN 26
Db 21 QVTDHLEFYLNVEMQWTTCKPRTTN 46
XX
RESULT 9
AAU79863
ID AAU79863 standard; peptide: 49 AA.
XX

```

```

AC AAU79863;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8) antigenic fragment #11.
XX
KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
XX spermatoogenesis; seminal fluid viscosity; cryopreserved sperm;
XX sperm motility; fertilisation; antigenic peptide.
XX
OS Homo sapiens.
XX
PN WO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Holloway JL, Gao Z, Bishop PD;
XX
DR WPI; 2002-383044/41.
XX
PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatoogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 97-98; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatoogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 49 AA;
XX
Query Match 78.8%; Score 149; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. NO. 5.6e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 QVTDHLEFYLNVEMQWTTCKPRTTN 26
Db 24 QVTDHLEFYLNVEMQWTTCKPRTTN 49
XX
RESULT 10
AAU79861
ID AAU79861 standard; peptide: 24 AA.
XX
AC AAU79861;
XX
DT 15-JUL-2002 (first entry)
XX

```

DE Human cystatin-8 (Zcys8) antigenic fragment #9.
 XX
 XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 PS Claim 2; Page 97; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterizing sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 SQ Sequence 24 AA;
 QY
 Query Match 70.4%; Score 133; DB 5; Length 24;
 Best Local Similarity 100.0%; Pred. No. 4e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 QVTDHLEHYHLNVEMQWTTCKPKE 23
 2 QVTDHLEHYHLNVEMQWTTCKPKE 24
 RESULT 11
 ID AAU79867 standard; peptide; 48 AA.
 AC AAU79867;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cystatin-8 (Zcys8) antigenic fragment #15.
 XX
 KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

KM sperm motility; fertilisation; antigenic peptide.
 XX
 XX Homo sapiens.
 XX
 PN WO200220567-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US026868.
 XX
 PR 01-SEP-2000; 2000US-0230230P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Holloway JL, Gao Z, Bishop PD;
 DR WPI; 2002-383044/41.
 XX
 PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.
 PS Claim 2; Page 99; 100pp; English.
 XX
 CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis, modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterizing sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 CC
 SQ Sequence 48 AA;
 QY
 Query Match 70.4%; Score 133; DB 5; Length 48;
 Best Local Similarity 100.0%; Pred. No. 8.5e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 NVEMQWTTCKPRTTNCVPOERE 33
 1 NVEMQWTTCKPRTTNCVPOERE 23
 RESULT 12
 ID ADD46708 standard; protein; 142 AA.
 AC ADD46708;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein AAC36317, SEQ ID NO 12393.
 XX
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KM chronic constriction injury; CCI, spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003016475-A2.

[illegible]

```

XX SS      Rattus norvegicus.
XX XX      WO2003016475-A2.
XX XX      27-FEB-2003.
XX XX      14-AUG-2002; 2002MO-US025765.
XX XX      14-AUG-2001; 2001US-0312147P.
XX XX      01-NOV-2001; 2001US-0346382P.
XX XX      26-NOV-2001; 2001US-0333347P.
XX XX      (GENO ) GEN HOSPITAL CORP.
XX XX      (FARB ) BAYER AG.
XX XX      Woolf C, D'urso D, Befort K, Costigan M;
XX XX      WPI; 2003-268312/26.
XX XX      GENBANK; AAC6317.
XX XX      New composition comprising two or more isolated polypeptides, useful for
XX XX      preparing a medicament for treating pain in an animal.
XX XX      Claim 1; Page: 1017pp; English.
XX XX      The invention discloses a composition comprising two or more isolated rat
XX XX      or human polynucleotides or a polynucleotide which represents a fragment,
XX XX      derivative or allelic variation of the nucleic acid sequence. Also
XX XX      claimed are a vector comprising the novel polynucleotide, a host cell
XX XX      comprising the vector, a method for identifying a nucleotide sequence
XX XX      which is differentially regulated in an animal subjected to pain and a
XX XX      kit to perform the method, an array, a method for identifying an agent
XX XX      that increases or decreases the expression of the polynucleotide sequence
XX XX      that is differentially expressed in neuronal tissue of a first animal
XX XX      subjected to pain, a method for identifying a compound which regulates
XX XX      the expression of a polynucleotide sequence which is differentially
XX XX      expressed in an animal subjected to pain, a method for identifying a
XX XX      compound that regulates the activity of one or more of the
XX XX      polynucleotides, a method for producing a pharmaceutical composition, a
XX XX      method for identifying a compound or small molecule that regulates the
XX XX      activity in an animal of one or more of the polypeptides given in the
XX XX      specification, a method for identifying a compound useful in treating
XX XX      pain and a pharmaceutical composition comprising the one or more
XX XX      polypeptides or their antibodies. The polynucleotide or the compound that
XX XX      modulates its activity is useful for preparing a medicament for treating
XX XX      pain (e.g. spinal segmental nerve injury (SNI) in an animal (e.g. gene
XX XX      injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
XX XX      therapy). The sequence presented is a rat protein (shown in Table 2 of
XX XX      the specification) which is differentially expressed during pain. Note:
XX XX      The sequence data for this patent did not form part of the printed
XX XX      specification, but was obtained in electronic form directly from WIPO at
XX XX      ftp.wipo.int/pub/published_pct_sequences.
XX XX      Sequence 142 AA:
XX XX
XX XX      Query Match          50.3%; Score 95; DB 7; Length 142;
XX XX      Best Local Similarity 45.5%; Pred. No. 4,6e-05;
XX XX      Matches 15; Conservative 10; Mismatches 6; Indels 2; Gaps 1;
XX XX
XX XX      QY      1 QVTDHLEHYHNVEMQWTTCKP--ETTNVCPQE 31
XX XX      [:::||||:::|::| |::|::|:
XX XX      Db      77 QITDMEYHIDVQISRSNGCRKPLNTNENCIPQK 109
XX XX
XX XX      RESULT 14
XX XX      ID      AAE02404 standard; protein: 142 AA.
XX XX      AC      AAE02404;
XX XX      XT      10-AUG-2001 (first entry)
XX XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11 ; Search time 8.44643 Seconds

(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189

Sequence: 1 QVTDHLEHYHNVEMQWTTCKPRTNVCVPOERE 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A COMB pep:*

2: /cgn2_6/ptodata/2/1aa/5B COMB pep:*

3: /cgn2_6/ptodata/2/1aa/6A COMB pep:*

4: /cgn2_6/ptodata/2/1aa/6B COMB pep:*

5: /cgn2_6/ptodata/2/1aa/PCTUS COMB pep:*

6: /cgn2_6/ptodata/2/1aa/backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	47.1	142	3	US-09-431-480-3
2	89	47.1	142	3	US-09-617-302-3
3	70	37.0	142	3	US-09-431-480-4
4	70	37.0	142	3	US-09-617-302-4
5	67	35.4	141	3	US-09-431-480-2
6	67	35.4	141	3	US-09-617-302-2
7	65	34.4	127	4	US-08-849-303-19
8	56	29.6	140	4	US-09-886-319A-46
9	56	29.6	140	4	US-09-886-319A-48
10	54.5	28.8	111	4	US-09-849-303-26
11	54	28.6	120	4	US-09-775-932-2
12	54	28.6	120	6	5432264-4
13	54	28.6	145	2	US-08-832-535-11
14	54	28.6	146	2	US-08-791-522-3
15	54	28.6	146	3	US-08-744-138-3
16	54	28.6	146	3	US-09-019-485-4
17	54	28.6	146	3	US-09-314-777-3
18	54	28.6	146	3	US-09-431-480-6
19	54	28.6	146	3	US-09-617-302-6
20	54	28.6	146	4	US-09-241-376-3
21	54	28.6	146	4	US-09-528-436B-3
22	54	28.6	146	4	US-09-886-319A-47
23	54	28.6	146	4	US-09-940-497-3
24	54	28.6	146	4	US-09-976-594-37
25	54	28.6	146	4	US-08-849-303-17
26	54	28.6	146	5	PCT-US95-07135-9
27	54	28.6	146	6	5432264-6

28	53	28.0	122	4	US-09-775-932-10	Sequence 10, Appl
29	53	28.0	142	3	US-08-744-138-4	Sequence 4, Appl
30	53	28.0	142	3	US-09-431-480-7	Sequence 7, Appl
31	53	28.0	142	3	US-09-617-302-7	Sequence 7, Appl
32	53	28.0	142	4	US-09-241-376-4	Sequence 4, Appl
33	53	28.0	142	4	US-09-940-497-4	Sequence 4, Appl
34	53	28.0	142	4	US-09-976-594-358	Sequence 358, App
35	53	28.0	142	4	US-08-849-303-20	Sequence 20, Appl
36	53	28.0	604	4	US-09-820-809-13	Sequence 13, Appl
37	53	28.0	912	5	PCT-US91-09422-19	Sequence 19, Appl
38	52	27.5	140	3	US-09-431-480-5	Sequence 5, Appl
39	52	27.5	140	3	US-09-617-302-5	Sequence 5, Appl
40	52	27.5	140	4	US-08-849-303-18	Sequence 18, Appl
41	50	26.5	68	4	US-08-858-207A-492	Sequence 492, App
42	50	26.5	112	4	US-08-849-303-16	Sequence 16, Appl
43	50	26.5	118	4	US-09-775-932-24	Sequence 24, Appl
44	50	26.5	632	1	US-08-295-814B-10	Sequence 10, Appl
45	50	26.5	632	3	US-09-343-361-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-09-431-480-3
Sequence 3, Application US/09431480
Patent No. 6235708
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72
CURRENT APPLICATION NUMBER: US/09/431,480
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 60/109,217
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/156,382
EARLIER FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 142
TYPE: PRT
ORGANISM: Mus musculus
US-09-431-480-3

Query Match      47.1%; Score 89; DB 3; Length 142;
Best Local Similarity 41.2%; Pred. No. 1.4e-05;
Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

QY      1 QVTDHLEHYHNVEMQWTTCKP--ETNVCVPOER 32
Db      77 QITDREMYQIDVOISRSNCKKPLNTENCIPQK 110

RESULT 2
US-09-617-302-3
Sequence 3, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
```

SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-617-302-3

Query Match 47.1%; Score 89; DB 3; Length 142;
Best Local Similarity 41.2%; Pred. No. 1.4e-05;
Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 32
Db 77 QITDMEYQIDVOISRSNCKKPLNTENCIPQK 110

RESULT 3
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match 37.0%; Score 70; DB 3; Length 142;
Best Local Similarity 48.5%; Pred. No. 0.0085;
Matches 16; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 31
Db 77 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 109

RESULT 4
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-4

Query Match 37.0%; Score 70; DB 3; Length 142;
Best Local Similarity 48.5%; Pred. No. 0.0085;
Matches 16; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 31
Db 77 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 109

RESULT 5
US-09-431-480-2
; Sequence 2, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-2

Query Match 35.4%; Score 67; DB 3; Length 141;
Best Local Similarity 42.4%; Pred. No. 0.023;
Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 31
Db 76 QITDSELYLEVNIAITMCKKIAGDNENCLFQ 108

RESULT 6
US-09-617-302-2
; Sequence 2, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-2

Query Match 35.4%; Score 67; DB 3; Length 141;
Best Local Similarity 42.4%; Pred. No. 0.023;
Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHLNVEMQWTTCKP--ETNVCVPOE 31
Db 76 QITDSELYLEVNIAITMCKKIAGDNENCLFQ 108

RESULT 7
US-08-849-303-19
; Sequence 19, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urvain, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,303
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-08-849-303-19

Query Match 34.4%; Score 65; DB 4; Length 127;
Best Local Similarity 48.3%; Pred. No. 0.04;
Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYLNVMQWTTCKPPT--TNC 27
Db 62 QLVAGVNYFLDVEWGRITTCSTQNLITNC 90

RESULT 8
US-09-886-319A-46
; Sequence 46, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081

PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-886-319A-46

Query Match 29.6%; Score 56; DB 4; Length 140;
Best Local Similarity 44.8%; Pred. No. 0.95;
Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYLNVMQWTTCKPPT--TNC 27
Db 75 QLVAGVNYFLDVEWGRITTCSTQNLITDC 103

RESULT 9
US-09-886-319A-48
; Sequence 48, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-886-319A-48

Query Match 29.6%; Score 56; DB 4; Length 140;
Best Local Similarity 44.8%; Pred. No. 0.95;
Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYLNVMQWTTCKPPT--TNC 27
Db 75 QLVAGVNYFLDVEWGRITTCSTQNLITDC 103

RESULT 10
US-08-849-303-26
; Sequence 26, Application US/08849303
; Patent No. 6680424
; GENERAL INFORMATION:
; APPLICANT: Atkinson, Howard J.
; APPLICANT: McPherson, Michael J.
; APPLICANT: Urvain, Peter E.
; TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey

COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-26

Query Match 28.6%; Score 54.5; DB 4; Length 111;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC-VPOERE 33
DB 47 QVVGCVKTYLMEELKTTCKTVKPRKGYQEIQNCNLPPEHQ 88

RESULT 11
US-09-775-932-2
Sequence 2, Application US/09775932
Patent No. 6534477
GENERAL INFORMATION:
APPLICANT: University of British Columbia
TITLE OF INVENTION: Production and use of Modified Cycatutins
FILE REFERENCE: 58069
CURRENT APPLICATION NUMBER: US/09/775,932
CURRENT FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: CA99/00717
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: 60/095,503
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 120
TYPE: PRT
ORGANISM: Homo sapiens
US-09-775-932-2

Query Match 28.6%; Score 54; DB 4; Length 120;
Best Local Similarity 41.4%; Pred. No. 1.6;
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27
DB 55 QIVAGVNYFLDVELGRITCTKTPNLDNC 83

RESULT 12
5432264-4
Patent No. 5432264
APPLICANT: GRUBB, ANDERS;LUNDWALL, AKE;ABRAHAMSON, MAGNUS;

DALBOGE, HENRIK
TITLE OF INVENTION: RECOMBINANT 3-DES-OH-CYSTATIN C PRODUCED
BY EXPRESSION IN A PROCARYOTIC HOST CELL
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/929,290
FILING DATE: 13-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 440,221
FILING DATE: 21-NOV-1989
APPLICATION NUMBER: 297,198
FILING DATE: 20-MAY-1988
SEQ ID NO:4
LENGTH: 120
5432264-4

Query Match 28.6%; Score 54; DB 6; Length 120;
Best Local Similarity 41.4%; Pred. No. 1.6;
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27
DB 55 QIVAGVNYFLDVELGRITCTKTPNLDNC 83

RESULT 13
US-08-832-535-11
Sequence 11, Application US/08832535
Patent No. 5919658
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: LI, HAODONG
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L
TITLE OF INVENTION: HUMAN CYSTATIN F
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832,535
FILING DATE: 03-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KIMBALL, PAUL C.
REGISTRATION NUMBER: 34,610
REFERENCE/DOCKET NUMBER: PF265
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1700
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-832-535-11

Query Match 28.6%; Score 54; DB 2; Length 145;
Best Local Similarity 41.4%; Pred. No. 1.9;
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27

Db 81 QIVAGVNYFLDVELGRTTCTKTPNLNDC 109

RESULT 14

US-08-791-522-3
; Sequence 3, Application US/08791522
; Patent No. 5935817
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791.522
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 181387
; US-08-791-522-3

Query Match 28.6%; Score 54; DB 2; Length 146;
Best Local Similarity 41.4%; Pred. No. 2;
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27
Db 81 QIVAGVNYFLDVELGRTTCTKTPNLNDC 109

RESULT 15

US-08-744-138-3
; Sequence 3, Application US/08744138
; Patent No. 6011012
; GENERAL INFORMATION:
; APPLICANT: Gentz, Retner L.
; APPLICANT: Ni, Jian
; APPLICANT: Rosen, Craig A.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: Human Cystatin E
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/744.138
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF202P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301 309 8504
; TELEFAX: 301 309 8512
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Cystatin C
; US-08-744-138-3

Query Match 28.6%; Score 54; DB 3; Length 146;
Best Local Similarity 41.4%; Pred. No. 2;
Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27
Db 81 QIVAGVNYFLDVELGRTTCTKTPNLNDC 109

Search completed: March 18, 2004, 14:25:37
Job time : 9.44643 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:15:06 ; Search time 19.9702 Seconds

(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189

Sequence: 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPORE 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	33	9	US-09-941-314-12 Sequence 12, Appl
2	189	100.0	52	9	US-09-941-314-14 Sequence 14, Appl
3	189	100.0	59	9	US-09-941-314-16 Sequence 16, Appl
4	189	100.0	80	9	US-09-941-314-15 Sequence 15, Appl
5	189	100.0	115	9	US-09-941-314-3 Sequence 3, Appl1
6	189	100.0	117	9	US-09-941-314-4 Sequence 4, Appl1
7	189	100.0	137	9	US-09-941-314-2 Sequence 2, Appl1
8	149	78.8	46	9	US-09-941-314-10 Sequence 10, Appl
9	149	78.8	49	9	US-09-941-314-13 Sequence 13, Appl
10	133	70.4	24	9	US-09-941-314-11 Sequence 11, Appl
11	133	70.4	48	9	US-09-941-314-17 Sequence 17, Appl
12	72.5	38.4	203	9	US-09-893-737-164 Sequence 164, App
13	68.5	36.2	138	14	US-10-239-663-37 Sequence 37, Appl
14	67.5	35.7	37	9	US-09-864-761-40986 Sequence 40986, A
15	65	34.4	127	8	US-08-849-303-19 Sequence 19, Appl

16	63	33.3	145	9	US-09-740-638-2	Sequence 2, Appl1
17	63	33.3	145	13	US-10-006-467-2	Sequence 2, Appl1
18	63	33.3	145	14	US-10-235-148-2	Sequence 14, Appl
19	63	33.3	145	14	US-10-168-425-14	Sequence 5, Appl1
20	63	33.3	165	9	US-09-740-638-5	Sequence 5, Appl1
21	63	33.3	165	13	US-10-006-467-5	Sequence 5, Appl1
22	63	33.3	165	13	US-10-235-148-5	Sequence 46, Appl
23	56	29.6	140	14	US-10-376-564-46	Sequence 48, Appl
24	56	29.6	140	14	US-10-376-564-48	Sequence 26, Appl1
25	55.5	29.4	633	10	US-09-911-458-2	Sequence 2, Appl1
26	55.5	29.4	633	14	US-10-298-974-2	Sequence 2, Appl1
27	54.5	28.8	111	8	US-08-849-303-26	Sequence 26, Appl1
28	54.5	28.8	120	9	US-09-775-932-2	Sequence 17, Appl
29	54	28.6	146	8	US-08-849-303-17	Sequence 3, Appl1
30	54	28.6	146	9	US-09-940-497-3	Sequence 3, Appl1
31	54	28.6	146	9	US-09-969-934-3	Sequence 3, Appl1
32	54	28.6	146	14	US-10-329-428-3	Sequence 47, Appl
33	54	28.6	146	14	US-10-376-564-47	Sequence 10, Appl
34	53	28.0	122	9	US-09-775-932-10	Sequence 20, Appl
35	53	28.0	142	8	US-08-849-303-20	Sequence 4, Appl1
36	53	28.0	142	9	US-09-940-497-4	Sequence 132, App
37	53	28.0	142	12	US-10-262-839-132	Sequence 13, Appl
38	53	28.0	604	9	US-09-820-809-13	Sequence 2, Appl1
39	53	28.0	983	14	US-10-255-149-2	Sequence 18, Appl
40	52	27.5	140	8	US-08-849-303-18	Sequence 61129, A
41	51.5	27.2	678	12	US-10-282-1228-61129	Sequence 191984, A
42	51	27.0	447	12	US-10-424-599-191984	Sequence 16, Appl
43	50	26.5	109	14	US-10-029-386-30645	Sequence 24, Appl
44	50	26.5	112	8	US-08-849-303-16	
45	50	26.5	118	9	US-09-775-932-24	

ALIGNMENTS

RESULT 1
US-09-941-314-12
Sequence 12, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941,314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-12

Query Match 100.0%; Score 189, DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPORE 33
DB 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPORE 33

RESULT 2
US-09-941-314-14
Sequence 14, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
Inhibit Cancer Procoagulant Protein

```
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14
/ LENGTH: 52
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-14
```

```
Query Match          100.0%; Score 189; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 1,7e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33
Db      20 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 52
```

```
RESULT 3
US-09-941-314-16
/ Sequence 16, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: Zymogenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 16
/ LENGTH: 59
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-16
```

```
Query Match          100.0%; Score 189; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 1,9e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33
Db      2 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 34
```

```
RESULT 4
US-09-941-314-15
/ Sequence 15, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: Zymogenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15
/ LENGTH: 80
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-15
```

```
Query Match          100.0%; Score 189; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 2,7e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33
Db      23 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 55
```

```
RESULT 5
US-09-941-314-3
/ Sequence 3, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: Zymogenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 115
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-3
```

```
Query Match          100.0%; Score 189; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 4e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33
Db      53 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 85
```

```
RESULT 6
US-09-941-314-4
/ Sequence 4, Application US/09941314
/ Patent No. US20020142396A1
/ GENERAL INFORMATION:
/ APPLICANT: Zymogenetics, Inc.
/ TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
/ TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
/ FILE REFERENCE: 00-81PC
/ CURRENT APPLICATION NUMBER: US/09/941.314
/ CURRENT FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/230,230
/ PRIOR FILING DATE: 2001-09-01
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 117
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-941-314-4
```

```
Query Match          100.0%; Score 189; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 4,1e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 33
Db      55 QVTDHLEYHLNVEMQWTTCKPFTTNCVPOERE 87
```

```
RESULT 7
US-09-941-314-2
/ Sequence 2, Application US/09941314
/ Patent No. US20020142396A1
```

```
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-2

Query Match      100.0%; Score 189; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 4,8e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 33
Db      75 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 107

RESULT 8
US-09-941-314-10
; Sequence 10, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-10

Query Match      78.8%; Score 149; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 4,4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 26
Db      21 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 46

RESULT 9
US-09-941-314-13
; Sequence 13, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 49
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-13

Query Match      78.8%; Score 149; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 4,7e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 26
Db      24 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 49

RESULT 10
US-09-941-314-11
; Sequence 11, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-11

Query Match      70.4%; Score 133; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 3,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 23
Db      2 QVTDHLEHYHLNVEQMWTTCCKPETTNCVPOERE 24

RESULT 11
US-09-941-314-17
; Sequence 17, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-17

Query Match      70.4%; Score 133; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 7,1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      11 NVEMQWTTTCCKPETTNCVPOERE 33
Db      1 NVEMQWTTTCCKPETTNCVPOERE 23
```

RESULT 12
US-09-893-737-164
; Sequence 164, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 164
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-164

Query Match 38.4%; Score 72.5; DB 9; Length 203;
Best Local Similarity 50.0%; Pred. No. 0.064; Indels 1; Gaps 1;
Matches 16; Conservative 5; Mismatches 10;

QY 1 QVTDLEHYHNVEMQWTTQCK-PEPTNCVPOE 31
DB 115 QVTDSEYIEYEVKIAFTICKKISEDCNCAFOE 146

RESULT 13
US-10-239-663-37
; Sequence 37, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kahnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GPO50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-37

Query Match 36.2%; Score 68.5; DB 14; Length 138;
Best Local Similarity 46.9%; Pred. No. 0.15; Indels 1; Gaps 1;
Matches 15; Conservative 6; Mismatches 10;

QY 1 QVTDLEHYHNVEMQWTTQCK-PEPTNCVPOE 31
DB 74 KVTDSEYIEYEVKIAFTICKKISEDCNCAFOE 105

RESULT 14
US-09-864-761-40986
; Sequence 40986, Application US/09864761

; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40986
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121894.14
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EST HUMAN HIT: AW665128.1, EVALUATE 1.10e-02
; OTHER INFORMATION: SWISSPROT HIT: O89869, EVALUATE 1.00e-04
US-09-864-761-40986

Query Match 35.7%; Score 67.5; DB 9; Length 37;
Best Local Similarity 48.4%; Pred. No. 0.05; Indels 1; Gaps 1;
Matches 15; Conservative 5; Mismatches 10;

QY 2 VTDLLEHYHNVEMQWTTQCK-PEPTNCVPOE 31
DB 1 VTDSLEYIEYEVKIAFTICKKISEDCNCAFOE 31

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 6.35119 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-12
Perfect score: 189
Sequence: 1 QVTDHLEHYHNLNEMQWTTCKPRTNCPQERE 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	47.1	139	2 A45361	cytastatin-related e
2	65	34.4	120	2 S10587	cytastatin C - rat
3	65	34.4	127	2 S07085	cytastatin C precurs
4	64	33.9	1779	2 T23130	hypothetical prote
5	55	29.1	4540	2 T30838	cytoplasmic dynein
6	54.5	28.8	111	2 A28793	cytastatin - putf ad
7	54	28.6	146	1 UDHU	cytastatin C precurs
8	54	28.6	302	2 B96520	hypothetical prote
9	54	28.6	463	2 E86464	hypothetical prote
10	53	28.0	142	2 A47142	cytastatin D precurs
11	53	28.0	912	2 JH0563	metabotropic gluta
12	53	28.0	3766	2 T29155	hypothetical prote
13	52	27.5	140	2 A36163	cytastatin C precurs
14	51.5	27.2	371	2 B96614	hypothetical prote
15	50	26.5	112	1 UDBO	cytastatin - bovine
16	50	26.5	232	2 T28369	probable uracil-DN
17	50	26.5	717	1 XNBVCM	glutamine-fructose
18	50	26.5	725	2 T42688	hypothetical prote
19	49.5	26.2	622	2 S63539	GABA/beta-alanine
20	49.5	26.2	953	2 T12577	H ₂ -exporting ATPas
21	49.5	26.2	1073	2 F89467	protein R09H3.1 [i
22	49	25.9	133	2 B86754	prophage p12 prote
23	49	25.9	328	2 S01359	salivary glue prot
24	49	25.9	457	2 C86454	hypothetical prote
25	49	25.9	695	2 D84634	hypothetical prote
26	49	25.9	1905	2 I51553	Plexin - African c
27	48.5	25.7	331	2 T41758	ACMNPV orf13 - Bom
28	48	25.4	133	2 JC4536	cytastatin precursor
29	48	25.4	455	2 T00856	pectate lyase (EC

30	48	25.4	603	2 T33134	hypothetical prote
31	47.5	25.1	214	1 A8BYD	ADP8 protein - Yea
32	47.5	25.1	1100	2 T47940	hypothetical prote
33	47.5	25.1	1275	2 T49362	hypothetical prote
34	47.5	25.1	1308	2 B75198	DNA helicase relat
35	47	24.9	180	2 G69222	molypdenum formylm
36	47	24.9	447	2 T19078	hypothetical prote
37	47	24.9	456	2 E84960	adenylosuccinate 1
38	47	24.9	555	2 S39953	isocitrate lyase (
39	47	24.9	610	2 T47725	hypothetical prote
40	47	24.9	1554	2 T06370	probable DNA (cyto
41	47	24.9	3161	2 T30342	protein HMWP1 - Ye
42	46.5	24.6	239	2 AB1834	pyridoxal phosphat
43	46.5	24.6	713	2 JC6012	glutamine-fructose
44	46.5	24.6	764	2 A84456	mutator-like trans
45	46	24.3	106	2 T27989	hypothetical prote

ALIGNMENTS

RESULT 1
A45361
cytastatin-related epididymal specific protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A45361
R:Corwall, G.A.; Orgebin-Crist, M.C.; Hann, S.R.
Mol. Endocrinol. 6, 1653-1664, 1992
A:Title: The CRE5 gene: a unique testis-regulated gene related to the cytastatin family is
A:Reference number: A45361; MUID:93078799; PMID:1280328
A:Accession: A45361
A:Status: preliminary; not compared with conceptual translation
A:Scatue: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-139 <COR>
A:Cross-reference: GB:849926; NID:9260492; PIDN:AAC35390.1; PID:9260493
A:Note: sequence extracted from NCBI backbone (NCBIPI:118813)
C:Superfamily: cytastatin; cytastatin homology
F:28-139/Domain: cytastatin homology <CTS>

Query Match 47.1%; Score 89; DB 2; Length 139;
Best Local Similarity 41.2%; Pred. No. 6.1e-05;
Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHNLNEMQWTTCKP--ETNVCVPOER 32
Db 74 QTFDMEYQIDVOISRNCKKPLNTENCIPOKK 107

RESULT 2
S10587
cytastatin C - rat
C:Species: Rattus sp. (rat)
C:Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C:Accession: S10587
R:Bernard, F.; Bernard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billard, M.; Gauch
Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990
A:Title: Rat cytastatin C: the complete amino acid sequence reveals a site for N-glycosyla
A:Reference number: S10587; MUID:90380276; PMID:2400577
A:Accession: S10587
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-120 <ESN>
A:Note: 43-Asn was also found
A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala
F:9-120/Domain: cytastatin homology <CVS>

Query Match 34.4%; Score 65; DB 2; Length 120;
Best Local Similarity 48.3%; Pred. No. 0.098;
Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy 1 QVTDHLEHYHNLNEMQWTTCKPRT--TNC 27

Db 55 QLVAGINYYLDVEMGRITTCSTQNTLINC 83

RESULT 3

S07085

cystatin C precursor - rat (fragment)

C/Species: Rattus norvegicus (Norway rat)

C/Date: 01-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999

C/Accession: S07085; S01337; S21109

R/Cole, T.; Dickson, P.W.; Esnard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Sch

Eur. J. Biochem. 186, 35-42, 1989

A/Title: The cDNA structure and expression analysis of the genes for the cysteine prote

A/Reference number: S07085; MUID:90092122; PMID:2689174

A/Accession: S07085

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-127 <COL>

A/Cross-references: EMBL:X16957; NID:956041; PIDN:CAA34831.1; PID:G736290

R/Esnard, A.; Esnard, F.; Faucher, D.; Gauthier, F.

FEBS Lett. 236, 475-478, 1988

A/Title: Two rat homologues of human cystatin C.

A/Reference number: S01337; MUID:88313020; PMID:3044831

A/Accession: S01337

A/Molecule type: protein

A/Residues: 8-49 <ESN>

R/Esnard, A.; Esnard, F.; Guillou, F.; Gauthier, F.

FEBS Lett. 300, 131-135, 1992

A/Title: Production of the cysteine proteinase inhibitor cystatin C by rat Sertoli cells

A/Reference number: S21109; MUID:92252121; PMID:1563513

A/Accession: S21109

A/Molecule type: protein

A/Residues: 8, 'XX', 11-20 <ES2>

C/Superfamily: cystatin; cystatin homology

A/Keywords: cysteine proteinase inhibitor

F/16-127/Dmain: cystatin homology <CTS>

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

F/80-90,104-124/Disulfide bonds: #status predicted

A/Map position: 1
A/Intons: 161/2; 223/2; 309/3; 332/2; 547/3; 603/1; 657/3; 745/2; 802/1; 856/1; 1031/1;
C/Superfamily: Caenorhabditis elegans hypothetical protein T28B8.3

Query Match 33.9%; Score 64; DB 2; Length 1779;
Best Local Similarity 34.2%; Pred. No. 2.1;
Matches 13; Conservative 8; Mismatches 7; Indels 10; Gaps 2;

Db 750 RITDHDYIINNPILKHPNVQMIT--DPTTHCLPK 785

RESULT 5

T30838

cytoplasmic dynein heavy chain - Parametium tetraurelia

C/Species: Parametium tetraurelia

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 02-Mar-2001

C/Accession: T30838; PC4341

R/Kandl, K.A.; Forney, J.D.; Asai, D.J.

submitted to the EMBL Data Library, January 1995

A/Description: The dynein genes of Parametium: the differential expression of axonemal ar

A/Reference number: Z20502

A/Accession: T30838

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4540 <KAN>

R/Asai, D.J.; Beckwith, S.M.; Kandl, K.A.; Keating, H.H.; Tjandra, H.; Forney, J.D.

J. Cell Sci. 107, 839-847, 1994

A/Title: The dynein genes of Parametium tetraurelia: Sequences adjacent to the catalytic

A/Reference number: PC4340; MUID:94334383; PMID:8056840

A/Accession: PC4341

A/Molecule type: mRNA

A/Residues: 1831-2029 <ASA>

C/Genetics: 57/3

A/Intons: 57/3

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

C/Superfamily: dynein heavy chain, cytosolic

Query Match 29.1%; Score 55; DB 2; Length 4540;
Best Local Similarity 40.6%; Pred. No. 94;
Matches 13; Conservative 5; Mismatches 8; Indels 6; Gaps 2;

Db 1191 ERYLNOQIGIEBQMT-SKPDSDCSPNEAE 1221

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

Query Match 28.8%; Score 54.5; DB 2; Length 111;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

Db 47 QVSGVITYLMBELTKTKTKTVGRPKGYQEIQNCNLPENQ 88

RESULT 7

UDHU

Cystatin C precursor (validated) - human

N/Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma C/Species: Homo sapiens (man)

C/Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000

C/Accession: S10216; S00004; J10095; A33400; S02751; A01270; A25434; S12288; A32732; A60

R/Abrahamson, M.; Olafsson, I.; Palseter, A.; Uysbaeck, M.; Lundwall, A.; Jensen, O.

Biochem. J. 268, 287-294, 1990

A/Title: Structure and expression of the human cystatin C gene.

A/Reference number: S10216; MUID:90303202; PMID:2363674

A/Accession: S10216

A/Molecule type: DNA

A/Residues: 1-146 <AB1>

A/Cross-references: EMBL:X52255; NID:930357; PIDN:CAA36497.1; PID:9296643

R/Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.

FEBS Lett. 216, 229-233, 1987

A/Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the

A/Reference number: S00004; MUID:87219149; PMID:3495457

A/Accession: S00004

A/Molecule type: mRNA

A/Residues: 1-146 <AB2>

A/Cross-references: EMBL:X05607; NID:930371; PIDN:CAA2906.1; PID:9755738

R/Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.

J. Exp. Med. 169, 1771-1778, 1989

A/Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a

A/Reference number: J10095; MUID:89235594; PMID:2541223

A/Accession: J10095

A/Molecule type: DNA

A/Residues: 1-146 <LEV>

A/Cross-references: GB:X61681; NID:930367; PIDN:CAA43856.2; PID:94490944

A/Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit

e)

R/Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada,

Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989

A/Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which

A/Reference number: A33400; MUID:89350949; PMID:2764935

A/Accession: A33400

A/Molecule type: DNA

A/Residues: 1-24, 'T', 26-146 <SAI>

A/Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:9181385; PIDN:AAA52164.1; PID:9

R/Ghiso, J.; Cowan, N.; Frangione, B.

Biol. Chem. Hoppe-Seyler 369, 205-208, 1988

A/Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron

A/Reference number: S02751; MUID:89076507; PMID:3264504

A/Accession: S02751

A/Molecule type: DNA

A/Residues: 82-119 <GH2>

A/Cross-references: EMBL:M27769

A/Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown

R/Grubb, A.; Lofberg, H.

Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982

A/Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in th

A/Reference number: A01270; MUID:92222268; PMID:6283552

A/Accession: A01270

A/Molecule type: protein

A/Residues: 27-131, 'S', 133-146 <GRU>

R/Ghiso, J.; Jensen, O.; Frangione, B.

Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986

A/Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland t

A/Reference number: A25434; MUID:86206076; PMID:3517880

A/Accession: A25434

A/Molecule type: protein

A/Residues: 37-93, 'Q', 95-146 <GHI>

R/Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Borchart, U.; Machleidt, W.

Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983

A/Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystati

A/Reference number: S01461; MUID:84110055; PMID:6662498

A/Accession: S12288

A/Molecule type: protein

A/Residues: 27-73 <TUR>

R/Brzin, J.; Popovic, T.; Turk, V.

Biochem. Biophys. Res. Commun. 118, 103-109, 1984

A/Title: Human cystatin, a new protein inhibitor of cysteine proteinases.

A/Reference number: A32732; MUID:84128015; PMID:6365094

A/Accession: A32732

A/Molecule type: protein

A/Residues: 27-76 <BER>

R/Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jensen, O.; Grubb, A.

Scand. J. Clin. Lab. Invest. 50, 85-93, 1990

A/Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst

A/Reference number: A60552; MUID:90193615; PMID:2315647

A/Accession: A60552

A/Molecule type: protein

A/Residues: 27-49, 'XX', 52-64 <OLA>

A/Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d

e defective gene is not present in CSF but is found instead in amyloid deposits

R/Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.

Biol. Chem. Hoppe-Seyler 371, 575-580, 1990

A/Title: Different forms of human cystatin C.

A/Reference number: S10607; MUID:91025625; PMID:2222856

A/Accession: S10607

A/Molecule type: protein

A/Residues: 27-53 <POP>

A/Experimental source: urine, kidney disease

A/Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also

R/Grubb, A.; Lofberg, H.; Barrett, A.J.

FEBS Lett. 170, 370-374, 1984

A/Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.

A/Reference number: S01462

A/Accession: S01462

A/Contents: annotation; disulfide bonds

R/Berti, P.J.; Storer, A.C.

Biochem. J. 302, 411-416, 1994

A/Title: Local pH-dependent conformational changes leading to proteolytic susceptibility

A/Reference number: S55305; MUID:94379969; PMID:8092991

A/Accession: S55305

A/Status: preliminary

A/Molecule type: protein

A/Residues: 27-49/106-146 <BER>

C/Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal fl

e patients with certain autoimmune diseases.

C/Comment: This protein is an inhibitor of cysteine proteinases and may serve an import

C/Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hemorrh

C/Genetics:

A/Gene: GDB:CST3

A/Cross-references: GDB:119817; OMIM:105150

A/Map position: 20p11.2-20p11.2

A/Intons: 81/3, 119/3

C/Superfamily: cystatin; cystatin homology

C/Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline

F.1.26/Domain: signal sequence #status predicted <SIG>

F.1.26/Product: cystatin C #status experimental <MAT>

F.35-146/Domain: cystatin homology <CTS>

F.81-85/Region: inhibitory #status predicted

F.29/Modified site: hydroxyproline (Pro) (partial) #status experimental

F.99-109,123-143/Disulfide bonds: #status experimental

Query Match 28.6%; Score 54; DB 1; Length 146;

Best Local Similarity 41.4%; Pred. No. 3.8;

Matches 12; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 QVTDHLEHYHNVEMQWTTCK--PETTNC 27

Db 81 QIVAGVYVFLDVELGRRTTKTQPNLNC 109

RESULT 8

B96520

hypochemical protein T2915.11 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: B96520

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B96520
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <STO>
 A:Cross-references: GB:AE005173; NID:g10645473; PIDN:AA621587.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T2U15.11
 A:Map position: 1

Query Match 28.6%; Score 54; DB 2; Length 302;
 Best Local Similarity 42.3%; Pred. No. 8;
 Matches 11; Conservative 4; Mismatches 7; Indels 4; Gaps 1;

OY 8 YHNVEMQWTTCKPFTTNCVPOERE 33
 Db 11 HHLDV----TSCPDATTCTDQDRK 32

RESULT 9
 E86464
 hypothetical protein F12G12.12 [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: E86464
 R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86464
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-463 <STO>
 A:Cross-references: GB:AE005172; NID:g10086471; PIDN:AA612531.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 28.6%; Score 54; DB 2; Length 463;
 Best Local Similarity 58.3%; Pred. No. 12;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 16 WTTCKPFTTNC 27
 Db 392 WVKCERPDFTNC 403

RESULT 10
 A47142
 Cystatin D precursor - human
 C:Species: *Homo sapiens* (man)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Jul-1999
 C:Accession: A47142; S18212
 R:Freije, J.P.; Balbin, M.; Abrahamson, M.; Velasco, G.; Dalboge, H.; Grubb, A.; Lopez-C J.; Biol. Chem. 269, 15737-15744, 1993
 A:Title: Human cystatin D. cDNA cloning, characterization of the *Escherichia coli* expres
 A:Reference number: A47142; MUID:93340179; PMID:8340398
 A:Accession: A47142
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-142 <PRE>

A:Cross-references: GB:X70377; NID:g398710; PIDN:CAA49838.1; PID:g398711
 A:Note: single residue difference between this report and S18218 was investigated and shc
 R:Freije, J.P.; Abrahamson, M.; Olafsson, I.; Velasco, G.; Grubb, A.; Lopez-Otin, C.
 J. Biol. Chem. 266, 20538-20543, 1991
 A:Title: Structure and expression of the gene encoding cystatin D, a novel human cysteine
 A:Reference number: S18212; MUID:92041895; PMID:1939105
 A:Accession: S18212
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-45, 'C', 47-142 <PRE>
 A:Cross-references: EMBL:X59964; NID:g30263; PIDN:CAA42590.1; PID:g30264
 C:Genetics:
 A:Gene: GDB:CST5
 A:Cross-references: GDB:136380; OMIM:123858
 A:Map position: 20p11.21-20p11.21
 A:Introns: 77/3; 115/3
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: cysteine proteinase inhibitor; extracellular protein; saliva
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:30-142/Domain: cystatin homology <CYS>

Query Match 28.0%; Score 53; DB 2; Length 142;
 Best Local Similarity 37.9%; Pred. No. 5;
 Matches 11; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

OY 1 QVTDHLEHNVEMQWTTCKP--PETTNC 27
 Db 77 QIVGVVYVFNVAKGRRTCTKSPNLDNC 105

RESULT 11
 JH0563
 metabotropic glutamate receptor 4 precursor - rat
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
 C:Accession: JH0563; I58149
 R:Tanabe, Y.; Mausu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors.
 A:Reference number: JH0561; MUID:92110002; PMID:1309649
 A:Accession: JH0563
 A:Molecule type: mRNA
 A:Residues: 1-912 <TRAN>
 A:Experimental source: Brain
 R:O'Hara, P.J.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.; Neuron 11, 41-52, 1993
 A:Title: The ligand-binding domain in metabotropic glutamate receptors is related to bact
 A:Reference number: I58149; MUID:9332699; PMID:8338667
 A:Accession: I58149
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-123, 'R', 125-912 <RES>
 A:Cross-references: GB:M90518; NID:g205400; PIDN:AA93190.1; PID:g205401
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions by me
 A:Gene: GLUR4
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phosphox
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>
 F:588-610/Domain: transmembrane #status predicted <TRI>
 F:625-645/Domain: transmembrane #status predicted <TRI>
 F:657-675/Domain: transmembrane #status predicted <TRI>
 F:700-720/Domain: transmembrane #status predicted <TRI>
 F:751-772/Domain: transmembrane #status predicted <TRI>
 F:786-807/Domain: transmembrane #status predicted <TRI>
 F:822-847/Domain: transmembrane #status predicted <TRI>
 F:98-301,454,484,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:621,669,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 28.0%; Score 53; DB 2; Length 912;
 Best Local Similarity 39.0%; Pred. No. 34;
 Matches 16; Conservative 3; Mismatches 6; Indels 16; Gaps 4;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 3.92857 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-12

Sequence: 1 QVTDHLEHYHNVEMQWTCCKPRTNCVPOERE 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	137	CS11_HUMAN	O9H112 homo sapien
2	104	55.0	139	CS11_MOUSE	O9D126 mus musculu
3	95	50.3	142	CS18_RAT	O88966 rattus norv
4	89	47.1	142	CS18_MOUSE	P22766 mus musculu
5	70	37.0	142	CS18_HUMAN	O60676 homo sapien
6	65	34.4	127	CYT8_RAT	P14841 rattus norv
7	63	33.3	165	CS1L_HUMAN	O9H114 homo sapien
8	57	30.2	146	CYT8_SAISC	O39093 haimtari sci
9	56	29.6	140	CYT8_MOUSE	P11460 mus musculu
10	55	29.1	4540	DYHC_PARTE	O27171 parametium
11	54.5	28.8	111	CYT8_BIRAR	P08933 bitis ariet
12	54	28.6	146	CYT8_HUMAN	P01034 homo sapien
13	54	28.6	146	CYT8_MACMU	O19092 macaca mula
14	53	28.0	142	CYT8_HUMAN	P28323 homo sapien
15	53	28.0	912	MGR4_RAT	P11423 rattus norv
16	52	27.5	148	CYT8_RABIT	O97862 cyttoclagus
17	50	26.5	148	CYT8_BOVIN	P01033 bos taurus
18	50	26.5	632	S6AB_HUMAN	P48066 homo sapien
19	50	26.5	611	Y252_DROME	O94065 drosophila
20	50	26.5	716	GFAL_YEAST	P14742 atabidopsia
21	50	26.5	801	PRD4_HUMAN	O9K055 homo sapien
22	49	25.9	107	TIM4_BOVIN	O97563 bos taurus
23	49	25.9	145	CYT8_HUMAN	O76096 homo sapien
24	49	25.9	170	TIM4_RABIT	O97591 cyttoclagus
25	49	25.9	224	TIM4_HUMAN	O99727 homo sapien
26	49	25.9	328	SGS3_DROER	P13730 drosophila
27	49	25.9	875	NPP3_HUMAN	O14638 haeconocle
28	49	25.9	1276	GIL1_ENTHI	P32022 entamoeba h
29	48.5	25.7	144	CYT8_MOUSE	O89056 mus musculu
30	48	25.4	455	PEL6_ARATH	O64510 arabidopsia
31	48	25.4	912	MGR4_HUMAN	O14833 homo sapien
32	47.5	25.1	214	PUR3_YEAST	P04161 eescharomyc
33	47	24.9	456	PUR3_BUCAI	P57351 buchiera ap

34	47	24.9	541	1	ACEA_YARLI	P41555 yarrowia li
35	46.5	24.6	239	1	PDJ1_ANASP	O82080 anabena sp
36	46.5	24.6	712	1	GFAL1_CANAL	P51704 candida alb
37	46.5	24.6	768	1	ENV1_SIVAI	P27757 simian immu
38	46	24.3	106	1	ILB1_CAEEL	O09626 caenorhabdi
39	46	24.3	367	1	HIS6_SYNY3	P73807 synechocyst
40	46	24.3	511	1	RNE_FORPU	P51211 porphyra pu
41	46	24.3	548	1	CUI2_SCHPO	O59755 schistosach
42	46	24.3	676	1	SYM_ECO57	O8X767 eescharichia
43	46	24.3	676	1	SYM_ECOL6	O8fx8 eescharichia
44	46	24.3	676	1	SYM_ECOL1	P00959 eescharichia
45	46	24.3	759	1	PEPX_STRGN	O93m42 streptococc

ALIGNMENTS

RESULT 1
ID CS11_HUMAN STANDARD; PRT; 137 AA.
AC O9H112; O9H113; 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin II precursor.
GN CS11 OR CS18L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leheslahti M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurtry A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Rose M.T., Scott C.B., Sehra H.K., Showkeen R., Sims S., Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sultoni J.B., Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:665-671(2001).
-1- SUBCELLULAR LOCATION: Secreted (Potential).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=O9H112-1; Sequence=Displayed;
Name=2;
IsoId=O9H112-2; Sequence=VSP_001260;
Note=No experimental confirmation available;
-1- SIMILARITY: Belongs to the cystatin family.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL096677; CAC13170.1; -
 DR EMBL; AL096677; CAC17423.1; -
 DR HSSP; P01038; 1A90.
 DR Genew; HGNC:15959; CST11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; Cy; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 DR Thiol protease inhibitor; Signal; Alternative splicing.
 DR SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 137 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT SITE 75 79 BY SIMILARITY.
 FT DISULFID 93 101 BY SIMILARITY.
 FT DISULFID 114 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 131 131 Missing (in isoform 2).
 FT VARSPLIC 76 110 /FtId=VSP_001260.
 SQ SEQUENCE 137 AA; 16375 MM; C58568C39A585C3B CRC64;

Query Match 100.0%; Score 189; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 1.7e-19;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33
 DB 75 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 107

RESULT 2
 CS11_MOUSE STANDARD; PRT; 139 AA.
 ID CS11_MOUSE
 AC Q9D269;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin 11 precursor.
 GN CST11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Epididymis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Rietshamn W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein W.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Togo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
 RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
 RA Hayashizaki Y.;
 RA Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AK020300; BAB32061.1; -
 DR HSSP; P01034; 1G96.
 DR MGD; MGI:1925490; Cst11.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; Cy; 1.
 DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
 DR Thiol protease inhibitor; Signal.
 DR SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 139 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT SITE 76 80 BY SIMILARITY.
 FT DISULFID 94 102 BY SIMILARITY.
 FT DISULFID 115 135 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 134 134
 SQ SEQUENCE 139 AA; 16217 MM; F228D9815FA32640 CRC64;

Query Match 55.0%; Score 104; DB 1; Length 139;
 Best Local Similarity 63.6%; Pred. No. 1.3e-07;
 Matches 21; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 33
 DB 76 QVTDHLEHYHNVEMQWTTCKPRTTNCVPOERE 108

RESULT 3
 CST8_RAT STANDARD; PRT; 142 AA.
 ID CST8_RAT
 AC O88969;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
 8).
 GN CST8 OR CRES.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;
 RX MEDLINE=99247899; PubMed=10259662;
 RA Cornwell G.A., Hsia N., Sutton H.G.;
 RA "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermatogenic gene";
 RL Biochem. J. 340:85-93(1999).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF090692; AAC36317.1; -
 DR HSSP; P01034; 1G96.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; Cy; 1.

KW Thiol protease inhibitor; Signal.
 PT SIGNAL 1 19 POTENTIAL.
 PT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT SITE 77 81 PROTEIN.
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 119 139 BY SIMILARITY.
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 142 AA; 16246 MW; FB873FAA6BCAB34 CRC64;
 Query Match 50.3%; Score 95; DB 1; Length 142; -
 Best Local Similarity 45.5%; Pred. No. 2,4e-06;
 Matches 15; Conservative 10; Mismatches 6; Indels 2; Gaps 1;
 QY 1 QVTDHLEHYLVNEMQWTTCKP--ETTNCVPOE 31
 77 QITDRMEHYDVOISRSNCKPLNTENCIPQK 109
 Db
 RESULT 4
 ID CSTR_MOUSE STANDARD; PRT; 142 AA.
 AC P32766; O89102;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin-
 related epididymal specific protein) (Cystatin 8).
 GN CSTR OR CRGS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH, and CD-1;
 RX MEDLINE=99247899; PubMed=10229662;
 RA Cornwall G.A., Hala N., Sutton H.G.;
 RT "Structure, alternative splicing and chromosomal localization of the
 cystatin-related epididymal spermatogenic gene.";
 RL Biochem. J. 340:85-93(1999).
 RN [2]
 RP SEQUENCE OF 4-142 FROM N.A.
 RC TISSUE=Epididymis;
 RX MEDLINE=93078799; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;
 RT "The CRGS gene: a unique testis-regulated gene related to the cystatin
 family is highly restricted in its expression to the proximal region
 of the mouse epididymis.";
 RL Mol. Endocrinol. 6:1653-1664(1992).
 RN [1]
 RP FUNCTION: Performs a specialized role during sperm development and
 maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 expression in the testis. Within the testis it is localized to the
 elongating spermatids, whereas within the epididymis it is
 exclusively synthesized by the proximal caput epithelium.
 CC -1- INDUCTION: Testicular factors or hormones other than androgens
 present in the testicular fluid may be involved in the regulation
 of CRGS gene expression.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF091503; AAC61754.1; -
 DR EMBL; AF090691; AAC6316.1; -
 DR EMBL; S49926; AAC35390.1; -

DR PIR; A4561; A4561.
 DR HSSP; P01034; 1G96.
 DR MCD; MG1:107161; Cstr8.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 KW Thiol protease inhibitor; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 FT SITE 77 81 PROTEIN.
 FT DISULFID 95 105 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 119 139 BY SIMILARITY.
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 4 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B98F6673E CRC64;
 Query Match 47.1%; Score 89; DB 1; Length 142;
 Best Local Similarity 41.2%; Pred. No. 1,7e-05;
 Matches 14; Conservative 11; Mismatches 7; Indels 2; Gaps 1;
 QY 1 QVTDHLEHYLVNEMQWTTCKP--ETTNCVPOE 32
 77 QITDRMEHYDVOISRSNCKPLNTENCIPQK 110
 Db
 RESULT 5
 ID CSTR_HUMAN STANDARD; PRT; 142 AA.
 AC O60676;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
 8).
 GN CSTR OR CRGS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=95344753; PubMed=7619504;
 RA Cornwall G.A., Hann S.R.;
 RT "Transient appearance of CRGS protein during spermatogenesis and
 caput epididymal sperm maturation.";
 RL Mol. Reprod. Dev. 41:37-46(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=1638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,
 RA Ellington J.A., Frankland J.A., Fraser J.A., French L., Garner P.,
 RA Graihm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.R., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lelevarasliho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McCormachie L.J., McHay K., McIneray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.D., Whitaker P., Willey D., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RL "The DNA sequence and comparative analysis of human chromosome 20.",
 CC Nature 414:865-871(2001).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 CC EMBL, AF059244; AAC14707.1; -
 CC EMBL, AL109954; CAB64234.1; -
 CC HSSP, P01034; 1G96.
 CC Genew, HGNC:2480; CST8.
 CC GO, GO:0004869; F:cysteine protease inhibitor activity; TAS.
 CC InterPro: IPR000010; Cystatin.
 CC Pfam, PF00031; cystatin; 1.
 CC SMART, SM00043; CY; 1.
 CC K101 protease inhibitor; Signal; Polymorphism.
 CC FT SIGNAL 1 21 POTENTIAL.
 CC FT CHAIN 22 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
 CC PROTEIN.
 CC FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 CC FT DISULFID 95 105 BY SIMILARITY.
 CC FT DISULFID 119 139 BY SIMILARITY.
 CC FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT VARIANT 142 142 A -> P (in dbSNP:1054633).
 CC /FTID=VAR_014527.
 CC FT SEQUENCE 142 AA; 16275 MW; 9A3512757E0F4ECD CRC64;
 CC SQ
 CC Query Match 37.0%; Score 70; DB 1; Length 142;
 CC Best Local Similarity 48.5%; Pred. No. 0.0076;
 CC Matches 16; Conservative 7; Mismatches 8; Indels 2; Gaps 1;
 CC
 CC QY 1 QVTDHLEVHNLNEMQWTCCKPRTN--CYPOE 31
 CC Db 77 QVTLNLEVLIDVEIARSDCKRPLSTNEICAIQE 109
 CC
 CC RESULT 6
 CC CYTC_RAT STANDARD; PRT; 127 AA.
 CC ID CYTC_RAT STANDARD; PRT; 127 AA.
 CC AC P14841;
 CC DT 01-APR-1990 (Rel. 14, Created)
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Cystatin C precursor (fragment).
 CC GN CST3.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OC NCBI_TaxID=10116;
 CC OX [1]
 CC RN SEQUENCE FROM N.A.
 CC RC STRAIN=Buffalo;
 CC RX MEDLINE=90092122; PubMed=2689174;
 CC RA Cole T., Dickson P.W., Esmard F., Averill F., Risbridger G.,
 CC Gauchier F., Schreiber G.,
 CC "The CDNA structure and expression analysis of the genes for the

RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
 RT in rat brain.";
 RL Eur. J. Biochem. 186:35-42(1989).
 RN [2]
 RN SEQUENCE OF 8-127.
 RX MEDLINE=90380276; PubMed=2400577;
 RA Benard F., Benard A., Faucher D., Capony J.-P., Derancourt J.,
 RA Brillard M., Gauchier F.,
 RT "Rat cystatin C: the complete amino acid sequence reveals a site for
 RT N-glycosylation.";
 RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
 RN [3]
 RN SEQUENCE OF 8-49.
 RX MEDLINE=88313020; PubMed=3044831;
 RA Benard A., Benard F., Faucher D., Gauchier F.,
 RT "Two rat homologues of human cystatin C.";
 RL FEBS Lett. 236:475-478(1988).
 RN [4]
 RN SEQUENCE OF 8-20.
 RP TISSUE=Sertoli cells;
 RC MEDLINE=92225121; PubMed=1563513;
 RX Benard A., Benard F., Guillon F., Gauchier F.,
 RT "Production of the cysteine proteinase inhibitor cystatin C by rat
 RT Sertoli cells.";
 RL FEBS Lett. 300:131-135(1992).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity. Known to inhibit cathepsin B,
 CC H, and L.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, X16957; CA314831.1; -
 CC PIR, S07085; S07085.
 CC PIR, S10587; S10587.
 CC HSSP, P01034; 1G96.
 CC InterPro: IPR000010; Cystatin.
 CC Pfam, PF00031; cystatin; 1.
 CC SMART, SM00043; CY; 1.
 CC DR PROSITE, PS00287; CYSTATIN; 1.
 CC K101 protease inhibitor; Signal.
 CC FT NON_TER 1 1
 CC FT SIGNAL 1 7
 CC FT CHAIN 8 127 CYSTATIN C.
 CC FT ACT SITE 18 18 REACTIVE SITE.
 CC FT SITE 62 66 SECONDARY AREA OF CONTACT.
 CC FT DISULFID 80 90 BY SIMILARITY.
 CC FT DISULFID 104 124 BY SIMILARITY.
 CC FT CONFLICT 25 25 A -> E (IN REF. 2).
 CC FT SEQUENCE 127 AA; 14039 MW; 78F70158B7925853 CRC64;
 CC SQ
 CC Query Match 34.4%; Score 65; DB 1; Length 127;
 CC Best Local Similarity 48.3%; Pred. No. 0.034;
 CC Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;
 CC
 CC QY 1 QVTDHLEVHNLNEMQWTCCKPRTN--TNC 27
 CC Db 62 QLVAGINLYIDVEIGRTTCKSQTNLTNC 90
 CC
 CC RESULT 7
 CC CSTL_HUMAN STANDARD; PRT; 165 AA.
 CC ID CSTL_HUMAN STANDARD; PRT; 165 AA.
 CC AC Q9H114;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=91054522; PubMed=2241983;
 RA Solem M., Rawson C., Lindburg K., Barnes D.;
 RT "Transforming growth factor beta regulates cystatin C in serum-free
 mouse embryo (SEME) cells."
 RL Biochem. Biophys. Res. Commun. 172:945-951(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv; TISSUE=liver;
 RX MEDLINE=95137392; PubMed=7835704;
 RA Huh C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
 RT "Structural organization, expression and chromosomal mapping of the
 mouse cystatin-C-coding gene (Cst3)."
 RL Gene 152:121-226(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ILS, and ISS;
 RX MEDLINE=21363810; PubMed=11471062;
 RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Camuff J.,
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Skeja J.M.;
 RT "High-throughput sequence identification of gene coding variants
 within alcohol-related QTLs."
 RL Mamm. Genome 12:657-663(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Straube R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Vallada D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Foley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S., Krzywinski M.I., Skalska J., Small D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59470; AAA63298.1; -
 CC EMBL; U10098; AAB41056.1; -
 CC EMBL; AF483486; AAL90760.1; -
 CC EMBL; AF483487; AAL90761.1; -
 CC EMBL; BC002072; AAH02072.1; -
 CC PIR; A36163; A36163.
 CC HSSP; P01034; 1G96.
 CC MGD; MGI:102519; Cst3.
 CC InterPro; IPR000010; Cystatin.
 CC Pfam; PF00031; Cystatin; 1.
 CC SMART; SM00043; Cy; 1.
 DR

DR PROSITE; PS00287; CYSTATIN; 1.
 KW Thiol protease inhibitor; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 140 CYSTATIN C.
 FT ACT SITE 31 31 REACTIVE SITE.
 FT SITE 75 79 SECONDARY AREA OF CONTACT.
 FT DISULFID 93 103 BY SIMILARITY.
 FT DISULFID 117 137 BY SIMILARITY.
 FT CONFLICT 16 16 A -> G (IN REF. 1).
 FT CONFLICT 84 84 L -> F (IN REF. 1).
 SQ SEQUENCE 140 AA; 15531 MW; 3A563406D58D0F5 CRC64;
 QY Query Match 29.6%; Score 56; DB 1; Length 140;
 DB Best Local Similarity 44.8%; Pred. No. 0.68;
 DB Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;
 QY 1 QVTDHLEHYHNVEMQWTTCKKPEF--TNC 27
 DB 75 QLVAGVNYFLDVEWGRITTKSQNTLTDC 103
 RESULT 10
 DYHC_PARTE STANDARD; PRT; 4540 AA.
 ID DYHC_PARTE
 AC Q27171;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dynein heavy chain, cytosolic (DYHC).
 GN DHC-8.
 OS Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 OC Paramcium.
 OX NCBI_TaxID=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Stock 51;
 RX MEDLINE=96157890; PubMed=8589455;
 RA Kandi K.A., Forney J.D., Asai D.J.;
 RT "The dynein genes of Paramecium tetraurelia: the structure and
 RT expression of the ciliary beta and cytoplasmic heavy chains."
 RL Mol. Biol. Cell 6:1549-1562(1995).
 CC -1- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
 CC retrograde motility of vesicles and organelles along microtubules.
 CC Dynein has ATPase activity; the force-producing power stroke is
 CC thought to occur on release of ADP.
 CC -1- SUBUNIT: Consists of at least two heavy chains and a number of
 CC intermediate and light chains.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
 CC (which binds cargo and interacts with other dynein components),
 CC and the head or motor domain. The motor contains six tandemly-
 CC linked AAA domains in the head, which form a ring. A stalk-like
 CC structure (formed by two of the coiled coil domains) protrudes
 CC between AAA 4 and AAA 5 and terminates in a microtubule-binding
 CC site. A seventh domain may also contribute to this ring; it is not
 CC clear whether the N-terminus or the C-terminus forms this extra
 CC domain. There are four well-conserved and two non-conserved ATPase
 CC sites, one per AAA domain. Probably only one of these (within AAA
 CC 1) actually hydrolyzes ATP, the others may serve a regulatory
 CC function.
 CC -1- SIMILARITY: Belongs to the dynein heavy chain family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U20449; AAA75445.1; -
 CC PIR; T30838; T30838.
 DR

```

DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR004273; Dyncin_heavy.
DR Pfam: PF03028; Dyncin_heavy; 1.
DR SMART: SM00382; AAA; 3.
KW Motor protein; Dyncin; Microtubule; ATP-binding; Repeat; Coiled coil.
FT DOMAIN 1 1796 STEM (BY SIMILARITY).
FT DOMAIN 1 1797 AAA 1 (BY SIMILARITY).
FT DOMAIN 2091 2348 AAA 2 (BY SIMILARITY).
FT DOMAIN 2457 2705 AAA 3 (BY SIMILARITY).
FT DOMAIN 2796 3056 AAA 4 (BY SIMILARITY).
FT DOMAIN 3076 3367 STALK (BY SIMILARITY).
FT DOMAIN 3444 3673 AAA 5 (BY SIMILARITY).
FT DOMAIN 3908 4123 AAA 6 (BY SIMILARITY).
FT DOMAIN 440 482 COILED COIL (POTENTIAL).
FT DOMAIN 698 722 COILED COIL (POTENTIAL).
FT DOMAIN 794 827 COILED COIL (POTENTIAL).
FT DOMAIN 975 995 COILED COIL (POTENTIAL).
FT DOMAIN 1169 1251 COILED COIL (POTENTIAL).
FT DOMAIN 1295 1311 COILED COIL (POTENTIAL).
FT DOMAIN 3076 3182 COILED COIL (POTENTIAL).
FT DOMAIN 3289 3367 COILED COIL (POTENTIAL).
FT DOMAIN 3653 3688 COILED COIL (POTENTIAL).
FT DOMAIN 3820 3851 COILED COIL (POTENTIAL).
FT DOMAIN 4238 4259 COILED COIL (POTENTIAL).
FT DOMAIN 4313 4342 COILED COIL (POTENTIAL).
FT NP_BIND 1835 1842 ATP (POTENTIAL).
FT NP_BIND 2129 2136 ATP (POTENTIAL).
FT NP_BIND 2496 2503 ATP (POTENTIAL).
FT NP_BIND 2834 2841 ATP (POTENTIAL).
SQ SEQUENCE 4540 AA; 528626 MW; 0C6103148BF85F9 CRC64;

Query Match 29.1%; Score 55; DB 1; Length 4540;
Best Local Similarity 40.6%; Pred. No. 34;
Matches 13; Conservative 5; Mismatches 8; Indels 6; Gaps 2;

7 EYHNL-----VEMQWTTCKPFTTCVPORE 33
Db 1191 ERYLNOQIOIEFQWMT-SKPDSDCSPNEAF 1221

RESULT 11
CYT BITAR STANDARD; PRT; 111 AA.
AC P08935;
AC 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB Cystatin.
OS Bitis arietans (African puff adder).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Bitis.
OX NCBI_Taxid=8692;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=88076861; PubMed=3500714;
RA Ritonja A., Evans H.J., Machleidt W., Barrett A.J.;
RT "Amino acid sequence of a cystatin from venom of the African puff
addor (Bitis arietans).";
RL Biochem. J. 246:799-802(1987).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the cystatin family.
DR HSSP; P01038; 1CEW.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00031; cystatin; 1.
DR SMART; SM00043; CY 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor.
FT ACT SITE 3 REACTIVE SITE.
FT SITE 47 51 SECONDARY AREA OF CONTACT.
FT DISUFID 65 81 PROBABLE.

```

```

FT VARIANT 57 57 M -> T (IN EQUAL AMOUNT).
SQ SEQUENCE 111 AA; 12678 MW; 1A31B6B246AC1A0C CRC64;

Query Match 28.8%; Score 54.5; DB 1; Length 111;
Best Local Similarity 33.3%; Pred. No. 0.87;
Matches 14; Conservative 8; Mismatches 11; Indels 9; Gaps 2;

1 QVTDHLEHNLVEMQWTTCKPFTTCVPORE 33
Db 47 QVSGVKRYLMEHLKTKTKVGRPKGYEIQNCLPBNQ 88

RESULT 12
CYTC_HUMAN STANDARD; PRT; 146 AA.
AC P01034;
AC 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
DE (Post-gamma-globulin).
GN CST3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=87219149; PubMed=3495457;
RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;
RT "Molecular cloning and sequence analysis of cDNA coding for the
precursor of the human cysteine proteinase inhibitor cystatin C.";
RL FEBS Lett. 216:229-233(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=90303202; PubMed=2363674;
RA Abrahamson M., Olafsson I., Palsdottir A., Ulvsbaeck M., Lundwall A.,
RA Jensen O., Grubb A.;
RT "Structure and expression of the human cystatin C gene.";
RL Biochem. J. 268:287-294(1990).
RN [3]
RP SEQUENCE FROM N.A. (HCHMA VARIANT).
RC TISSUE=Brain;
RX MEDLINE=89235594; PubMed=2541223;
RA Levy E., Lopez-Otin C., Ghiso J., Gellner D., Frangione B.;
RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is
related to a mutation in the cystatin C gene, an inhibitor of
cysteine proteases.";
RL J. Exp. Med. 169:1771-1778(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89350949; PubMed=2764935;
RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.;
RA Iemura S., Sanada K.;
RT "The human cystatin C gene (CST3) is a member of the cystatin gene
family which is localized on chromosome 20.";
RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
RN [5]
RP SEQUENCE FROM N.A.
RA Dickinson D.P., Hewett-Emmett D., Thiasse M.;
RT "Acquisition of complex patterns of differential expression in
epithelial cell populations during the evolution of type 2 cystatin
genes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,

```



```

DR EMBL; X59964; CAA42590.1; -.
DR EMBL; X70377; CAA49838.1; -.
DR EMBL; AL591074; CAC94785.1; -.
DR PIR; A47142; A47142.
DR HSSP; P01034; 1096.
DR Genew; HGNC:2477; CSTS.
DR MIM; 123858; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SMO0043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
DR Thiol protease inhibitor; Signal; Polymorphism.
KW SIGNAL
FT CHAIN 1 20
FT ACT SITE 21 142 CYSSTATIN D.
FT SITE 22 22 REACTIVE SITE (BY SIMILARITY).
FT DISULFID 95 105 SECONDARY AREA OF CONTACT.
FT DISULFID 119 139 BY SIMILARITY.
FT VARIANT 46 46 BY SIMILARITY.
                          C -> R (IN 45% OF THE POPULATION;
                          dBSNP:1799841).
                          /FTid=VAR 002208.
SQ SEQUENCE 142 AA; 16080 MM; CEFPA89BA87A0DA68 CRC64;
Query Match 28.0%; Score 53; DB 1; Length 142;
Best Local Similarity 37.9%; Pred. No. 1.8;
Matches 11; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 1 QVTDHLEHNVEMQWTTCK--PETNC 27
Db 77 QIVGVNVYFNKRGRTCTKSPNDNC 105

RESULT 15
MGR4_RAT STANDARD; PRT; 912 AA.
ID P31423;
AC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Metabotropic glutamate receptor 4 precursor (MGLUR4).
GN GMR4 OR GPRC4 OR MGLUR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92110002; PubMed=1309649;
RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
RT "A family of metabotropic glutamate receptors.";
RL Neuron 11:41-52(1993).
RN [3]
RP INTERACTION WITH PRKCAB.
RX MEDLINE=20571397; PubMed=11122333;
RA El Far O., Alras J., Wischmeyer E., Nehring R.B., Karschin A.,
RA Betz H.;
RT "Interaction of the C-terminal tail region of the metabotropic
RT glutamate receptor 7 with the protein kinase C substrate PICK1.";
RL Eur. J. Neurosci. 12:4215-4221(2000).
CC -1- FUNCTION: Receptor for glutamate. The activity of this receptor
CC is mediated by a G-protein that inhibits adenylate cyclase
CC activity.

```

```

CC -1- SUBUNIT: Interacts with PRKCAB.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Is widely distributed in the CNS. Predominant
CC expression is seen in the granule cells of the cerebellum.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR6.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92077; -; NOT ANNOTATED_CDS.
DR EMBL; M90518; AA93150.1; -.
DR PIR; JH0563; JH0563.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G-PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G-PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G-PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS50259; G-PROTEIN_RECP_F3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 32
FT CHAIN 33 912
FT DOMAIN 33 587 METABOTROPIC GLUTAMATE RECEPTOR 4.
FT TRANSSEM 588 610 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 611 624 I (POTENTIAL).
FT TRANSSEM 625 645 II (POTENTIAL).
FT DOMAIN 646 656 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 657 675 III (POTENTIAL).
FT DOMAIN 676 699 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 700 720 IV (POTENTIAL).
FT DOMAIN 721 750 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 751 772 V (POTENTIAL).
FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 786 808 VI (POTENTIAL).
FT DOMAIN 809 821 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 822 847 VII (POTENTIAL).
FT DOMAIN 848 912 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 124 124 Q -> R (IN REF. 2).
SQ SEQUENCE 912 AA; 101818 MM; 336430EPIB4B577 CRC64;
Query Match 28.0%; Score 53; DB 1; Length 912;
Best Local Similarity 39.0%; Pred. No. 12;
Matches 16; Conservative 3; Mismatches 6; Indels 16; Gaps 4;

QY 3 TDHLEHNVEMQWTTCK--PETNCVQERE 33
Db 496 TDHL-HLRTERMQPSSGOOLPRISICLP---CQGERK 530

```

Search completed: March 18, 2004, 14:16:12
Job time : 3.92857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 19.7738 Seconds
(without alignments)
526.560 Million cell updates/sec

Title: US-09-941-314-12

Perfect score: 189
Sequence: 1 QVTDHLEHYHNVEMQWTTCKPPTNCVPOERE 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriopl:*
- 17: sp_archaeopl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	189	100.0	138	4	Q8WXU6
2	95	50.3	139	11	Q8K5A3
3	67	35.4	141	11	Q9DAP1
4	67	35.4	141	11	Q80ZM5
5	64	33.9	1779	5	Q18150
6	57	30.2	540	5	Q8T9G8
7	57	30.2	723	9	Q94M04
8	57	30.2	778	4	Q8TAA0
9	57	30.2	966	5	Q8T9K4
10	57	30.2	966	5	Q9VLM8
11	57	30.2	966	5	Q9USB4
12	56	29.6	140	11	Q9BXP9
13	55.5	29.4	633	5	Q9US549
14	54	28.6	315	10	Q9C7Z0
15	54	28.6	315	10	Q8LAU8
16	54	28.6	463	10	Q93Z38

17	54	28.6	463	10	Q9FX14	Q9FX14 arabidopsis
18	53.5	28.3	634	5	Q76188	Q76188 manduca sex
19	53	28.0	367	5	Q7YTM6	Q7YTM6 caenorhabdi
20	53	28.0	400	13	Q8UVR3	Q8UVR3 xenopus lae
21	53	28.0	983	11	Q62916	Q62916 rattus norv
22	52	27.5	3766	5	Q17551	Q17551 caenorhabdi
23	52	27.5	236	16	Q8DS42	Q8DS42 streptococ
24	52	27.5	441	5	Q962Y5	Q962Y5 aecaris suu
25	52	27.5	1347	12	Q8VJ78	Q8VJ78 chlorella v
26	51.5	27.2	367	10	Q9CGR6	Q9CGR6 arabidopsis
27	51.5	27.2	371	10	Q9C6F7	Q9C6F7 arabidopsis
28	51.5	27.2	678	2	Q83XK8	Q83XK8 legionella
29	51.5	27.2	1458	10	Q8M367	Q8M367 oryza sativ
30	51.5	27.2	1458	10	Q7XD96	Q7XD96 oryza sativ
31	51.5	27.2	2165	5	Q8T1P5	Q8T1P5 dictyosteli
32	51	27.0	270	16	Q7V796	Q7V796 prochloroco
33	51	27.0	392	5	Q9VY83	Q9VY83 drosophila
34	51	27.0	591	16	Q8P9T0	Q8P9T0 leptospira
35	50.5	26.7	290	5	Q8SX83	Q8SX83 mus musculu
36	50.5	26.7	645	11	Q8CGW5	Q8CGW5 mus musculu
37	50	26.5	208	4	Q81YC9	Q81YC9 homo sapien
38	50	26.5	232	12	Q9YVN4	Q9YVN4 melanoplus
39	50	26.5	261	5	Q17476	Q17476 hydraclonia
40	50	26.5	711	4	Q81ZM6	Q81ZM6 homo sapien
41	50	26.5	711	4	Q81U92	Q81U92 homo sapien
42	50	26.5	891	5	Q81B61	Q81B61 plasmodium
43	49.5	26.2	256	9	Q7Y4K3	Q7Y4K3 streptococ
44	49.5	26.2	622	13	Q91503	Q91503 torpeda mar
45	49.5	26.2	953	10	P93Z65	P93Z65 mesembryant

ALIGNMENTS

RESULT 1

Q8WXU6 Q8WXU6 PRELIMINARY; PRT; 138 AA.
AC Q8WXU6:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE SC13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335480; ALU1991.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1
SO SEQUENCE 138 AA; 16506 MW; B49440ACA3585C64 CRC64;

Query Match 100.0%; Score 189; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.8e-20;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTTCKPPTNCVPOERE 33
Db 76 QVTDHLEHYHNVEMQWTTCKPPTNCVPOERE 108

RESULT 2

Q8K5A3 Q8K5A3 PRELIMINARY; PRT; 139 AA.
AC Q8K5A3:
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE Cystatin 11.
 GN CSF11.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Hamil K.G., Hall S.H.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF501290; AAM21709.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 139 AA; 1686 MW; E1E36DB786B4D08C CRC64;

Query Match 50.3%; Score 95; DB 11; Length 139;
 Best Local Similarity 54.5%; Pred. No. 3.7e-06;
 Matches 18; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 QVTDHLEHYLVNEMQWTTCK--PETTCVPORE 33
 DB 76 QMTNMFHITVEMQRTTCTKTKNLCNVQGE 108

RESULT 3
 Q9DAP1 PRELIMINARY; PRT; 141 AA.
 AC Q9DAP1;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE 1700006C19RIK protein.
 GN 1700006C19RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
 RA Wysshaw-Borja A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RL "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK005665; BAB24175.1; -
 DR HSSP; P01038; ICEW.
 DR KMD; MGI:1916544; 1700006C19RIK.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 141 AA; 16811 MW; C20FA0DB8B1AC378C CRC64;

Query Match 35.4%; Score 67; DB 11; Length 141;
 Best Local Similarity 42.4%; Pred. No. 0.051;
 Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYLVNEMQWTTCK--PETTCVPORE 31
 DB 76 QITDSLEYLVNEMQRTTCKKVGADNENCLFQ 108

RESULT 4

ID Q80ZNS PRELIMINARY; PRT; 141 AA.
 AC Q80ZNS;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE RIKEN cDNA 1700006C19 gene.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testicle;
 RA Strauberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC048681; AAH48681.1; -
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR InterPro; IPR003243; Cystatin_C/M.
 DR Pfam; PF00031; Cystatin; 1.
 DR ProDom; PD001231; Cystatin_C/M; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 141 AA; 16825 MW; C20FA0DB8B1AC378C CRC64;

Query Match 35.4%; Score 67; DB 11; Length 141;
 Best Local Similarity 42.4%; Pred. No. 0.051;
 Matches 14; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 QVTDHLEHYLVNEMQWTTCK--PETTCVPORE 31
 DB 76 QITDSLEYLVNEMQRTTCKKVGADNENCLFQ 108

RESULT 5

ID Q18150 PRELIMINARY; PRT; 1779 AA.
 AC Q18150; Q45627;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE T28B8.4 protein.
 GN T28B8.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA White S.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawke T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopia A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sultson J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinlock L., Wilkinson-Sproat J., Woldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81133; CAB03445.1; -
DR EMBL; AL021066; CAB03445.1; JOINED.
DR EMBL; AL021066; CA15925.1; -
DR EMBL; Z81133; CA15925.1; JOINED.
DR PIR; T23130; T23130.
DR WormPep; T2888.4; CE16519.
DR InterPro; IPR008938; ANM.
SQ SEQUENCE 1779 AA; 205511 MW; 70A4489148B96F3E CRC64;

Query Match 33.9%; Score 64; DB 5; Length 1779;
Best Local Similarity 34.2%; Pred. No. 1.9;
Matches 13; Conservative 8; Mismatches 7; Indels 10; Gaps 2;

Qy 1 QVTDHLEVHNLN-----VEMQWTCCKPPTTCVQ 30
Db 750 RITDHLDIINNLPTLKHPTNVQWIT--DPTPHCLPK 785

RESULT 6
Q8T9G8 PRELIMINARY; PRT; 540 AA.
AC Q8T9G8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE S001519P.
GN AATS-ALA OR CG13391.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Broketein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Paclib J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Gelinker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069773; AAL39918.1; -
DR FLYBase; FBgn0027094; Aats-ala.
DR GO; GO:0004813; F:alanine-tryptophan ligase activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006419; P:alanine-tryptophan aminoacylation; IEA.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR002318; RNA-synt_2c.
DR InterPro; IPR006193; RNA-synt_Ala.
DR Pfam; PF02272; DHHA1; 1.
DR Pfam; PF01411; RNA-synt_2c; 1.
DR PROSITE; PS50860; AA TRNA LIGASE II ALA; 1.
SQ SEQUENCE 540 AA; 59580 MW; B870DF9C386843B CRC64;

Query Match 30.2%; Score 57; DB 5; Length 540;
Best Local Similarity 35.7%; Pred. No. 6;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 1 QVTDHLEVHNLNEMQWTCCKPPTTCVQ 28
Db 157 KVGDELHLDIVERRMLTKNHSATHAL 184

RESULT 7

Q94M04 PRELIMINARY; PRT; 723 AA.
AC Q94M04;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Major core structural protein Pl.
GN 1.
OS Bacteriophage phi-12.
OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
OX NCBI_TaxID=161736;
RN [1]
RP SEQUENCE FROM N.A.
RA Gactieb P.J., Potgieter C., Wei H., Toporovsky I.;
RT "Characterization of Bacteriophage phi12.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF408636; AAL01109.1; -
DR InterPro; IPR006162; Phantne.S.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
SQ SEQUENCE 723 AA; 78950 MW; F5DEF7FF93A50B7C CRC64;

Query Match 30.2%; Score 57; DB 9; Length 723;
Best Local Similarity 50.0%; Pred. No. 8.1;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 HLEVHNLNEMQWTCCKP 22
Db 463 HTEYRFRVDLQWATCYLP 480

RESULT 8
Q8TA0 PRELIMINARY; PRT; 778 AA.
AC Q8TA0;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to PRAM-1 protein, PML-RARA target gene encoding an adaptor
DE molecule-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Testis;
RA Strauberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026171; AAL26171.1; -
SQ SEQUENCE 778 AA; 89091 MW; 405B01953B2704DD CRC64;

Query Match 30.2%; Score 57; DB 4; Length 778;
Best Local Similarity 38.7%; Pred. No. 8.7;
Matches 12; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

Qy 1 QVTDHLEVHNLNEMQWTCCKPPTTCVPOE 31
Db 148 QVTLGHLDPRELEDAWACETQETTEVPT 178

RESULT 9
Q8T9K4 PRELIMINARY; PRT; 966 AA.
AC Q8T9K4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE GMD3058P.
GN AATS-ALA OR CG13391.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```

OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Broketein P., Hong L., Agbayan A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celinker S.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY069255; ALU39400.1; -
DR FlyBase: FBgn0027094; Aats-ala.
DR GO: GO:0004813; F:alanine-tRNA ligase activity; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
DR InterPro: IPR003156; DHHA1.
DR InterPro: IPR002318; tRNA-synt_2c.
DR InterPro: IPR006193; tRNA_synt_Ala.
DR Pfam: PF02272; DHHA1; 1.
DR Pfam: PF01411; tRNA-synt_2c; 1.
DR PRINTS: PRO0980; TRNASYNTHALA.
DR TIGRFAMs: TIGR00344; alas; 1.
DR PROSITE: PSS0860; AA TRNA LIGASE II ALA; 1.
SQ SEQUENCE 966 AA; 107697 MW; 9F2079A119F463B4 CRC64;

Query Match 30.2%; Score 57; DB 5; Length 966;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTCQKPEPTNCV 28
DB 583 KVGDLELHIDVERRWLTMKHNSATHAL 610

Q9VLM8 PRELIMINARY; PRT; 966 AA.
ID Q9VLM8
AC Q9VLM8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE AATS-ALA protein.
GN AATS-ALA OR CG13391.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA Borczyk D., Buchanan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burdick K.C., Butam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Fouts D.E., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
RA Jaitani M., Kalish F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
RA Palazzolo M., Peltman G.S., Fan S., Pollard J., Furl V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003621; AAF52657.1; -
DR FlyBase: FBgn0027094; Aats-ala.
DR GO: GO:0004813; F:alanine-tRNA ligase activity; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
DR InterPro: IPR003156; DHHA1.
DR InterPro: IPR002318; tRNA-synt_2c.
DR InterPro: IPR006193; tRNA_synt_Ala.
DR Pfam: PF02272; DHHA1; 1.
DR Pfam: PF01411; tRNA-synt_2c; 1.
DR PRINTS: PRO0980; TRNASYNTHALA.
DR TIGRFAMs: TIGR00344; alas; 1.
DR PROSITE: PSS0860; AA TRNA LIGASE II ALA; 1.
SQ SEQUENCE 966 AA; 107741 MW; B1B63A7A8B87B6E CRC64;

Query Match 30.2%; Score 57; DB 5; Length 966;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QVTDHLEHYHNVEMQWTCQKPEPTNCV 28
DB 583 KVGDLELHIDVERRWLTMKHNSATHAL 610

Q9U6B4 PRELIMINARY; PRT; 966 AA.
ID Q9U6B4
AC Q9U6B4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alanyl-tRNA synthetase.
GN AATS-ALA OR ALAS OR CG13391.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Chihade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.,
RT "Detection of an intermediate Stage of Mitochondrial Genesis."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF188718; AAF05593.1; -
DR FlyBase: FBgn0027094; Aats-ala.
DR GO: GO:0004813; F:alanine-tRNA ligase activity; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006419; P:alanyl-tRNA aminoacylation; IEA.
DR InterPro: IPR003156; DHHA1.
DR InterPro: IPR002318; tRNA-synt_2c.
DR InterPro: IPR006193; tRNA_synt_Ala.
DR Pfam: PF02272; DHHA1; 1.
DR Pfam: PF01411; tRNA-synt_2c; 1.
DR PRINTS: PRO0980; TRNASYNTHALA.

```

DR TIGR00344; alas; 1.
 DR PROSITE; PSS0860; AA_TRNA_LIGASE_II_ALA; 1.
 KW Antisense tRNA synthetase.
 SQ SEQUENCE 966 AA; 107877 MW; 82034A189F0C81B5 CRC64;
 Query Match 30.2%; Score 57; DB 5; Length 966;
 Best Local Similarity 35.7%; Pred. No. 11;
 Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
 QY 1 QVTDHLEHYLVNEMQWTTCKPPTNCV 28
 DB 583 KVGDELEHIDVERHMTKMSATHAL 610
 RESULT 12
 Q9EPX9 PRELIMINARY; PRT; 140 AA.
 AC Q9EPX9; 01-MAR-2001 (TIGR00344, 16, Created)
 DT 01-MAR-2001 (TIGR00344, 16, Last sequence update)
 DT 01-JUN-2003 (TIGR00344, 24, Last annotation update)
 DE Cystatin C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C;
 RX MEDLINE=21010502; PubMed=11144350;
 RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakansson K., Grubb A.,
 "RGF-2-responsive neural stem cell proliferation requires CCG, a novel
 RT autocrine/paracrine cofactor.";
 RT Neuron 28:385-397(2000).
 RL EMBL; AF311741; AAC40283.1; -
 DR HSSP; P01034; 1G96.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 FT CHAIN 21 140
 FT VARIANT 16 A->G.
 FT VARIANT 84 L->F.
 SQ SEQUENCE 140 AA; 15517 MW; 3A563406DD58D785 CRC64;
 Query Match 29.6%; Score 56; DB 11; Length 140;
 Best Local Similarity 44.8%; Pred. No. 2.1;
 Matches 13; Conservative 5; Mismatches 9; Indels 2; Gaps 1;
 QY 1 QVTDHLEHYLVNEMQWTTCKPPTNCV 27
 DB 75 QLVAGVNYFLDVEWGRITTKSGTNLTDC 103
 RESULT 13
 Q9U5A9 PRELIMINARY; PRT; 633 AA.
 AC Q9U5A9; 01-MAY-2000 (TIGR00344, 13, Created)
 DT 01-MAY-2000 (TIGR00344, 13, Last sequence update)
 DT 01-JUN-2003 (TIGR00344, 24, Last annotation update)
 DE Amino acid transporter/amino acid-gated channel for sodium/potassium
 DE ions.
 GN CAATCH1.
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Dityptia; Sphingidae;
 OC Sphingidae; Sphinginae; Manduca.
 NCBI_Taxid=7130;
 [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Midgut;
 RX MEDLINE=20387341; PubMed=10829035;
 RA Feldman D.H., Harvey W.R., Stevens B.R.;
 RT "A novel electrogenic amino acid transporter is activated by K+ or
 RT Na+, is alkaline pH-dependent, and is Cl--independent.";
 RL J. Biol. Chem. 275:24518-24526(2000).
 DR EMBL; AF013963; AAF18560.1; -
 DR GO; GO:0005087; C:integral to plasma membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005328; F:neurotransmitter:sodium symporter activity; IEA.
 DR GO; GO:0006836; P:neurotransmitter transport; IEA.
 DR InterPro; IPR000175; Na/ntra_symport.
 DR Pfam; PF00209; SNF; 1.
 DR PRINTS; PR00176; NANEUSMPORT.
 DR ProDom; PD000446; Na/ntra_symport; 1.
 DR PROSITE; PSS0267; NA_NEUROTRAN_SYM_3; 1.
 SQ SEQUENCE 633 AA; 69934 MW; 597EDC58E98FDFD CRC64;
 Query Match 29.4%; Score 55.5; DB 5; Length 633;
 Best Local Similarity 42.9%; Pred. No. 12;
 Matches 12; Conservative 3; Mismatches 8; Indels 5; Gaps 2;
 QY 8 YHNVEMQWTTCKPPTNCVPOB 31
 DB 156 YHNVEMQWTTCKPPTNCVPOB 182
 RESULT 14
 Q9C7Z0 PRELIMINARY; PRT; 302 AA.
 AC Q9C7Z0;
 DT 01-JUN-2001 (TIGR00344, 17, Created)
 DT 01-JUN-2001 (TIGR00344, 17, Last sequence update)
 DT 01-JUN-2003 (TIGR00344, 24, Last annotation update)
 DE Desiccation-related protein, putative.
 GN T2J15.11.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Euphorbiales; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_Taxid=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gali J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Millechner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.L.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaynsberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.M.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 DR EMBL; AC051631; AAG51530.1; -
 DR PIR; B96520; B96520.
 SQ SEQUENCE 302 AA; 32808 MW; ED2B80339ECB64E4 CRC64;
 Query Match 28.6%; Score 54; DB 10; Length 302;
 Best Local Similarity 42.3%; Pred. No. 9.2;
 Matches 11; Conservative 4; Mismatches 7; Indels 4; Gaps 1;
 QY 8 YHNVEMQWTTCKPPTNCVPOB 33

DB 11 HHLQV---TSCPDQATTNCTDQDRK 32

RESULT 15

Q8LAU8 PRELIMINARY; PRT; 315 AA.
 AC Q8LAU8;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Desiccation-related protein, putative.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.U., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY087598; AM65140.1; -; SAFPCA9D6391F715 CRC64;
 SQ SEQUENCE 315 AA; 34328 MW; 5AFPCA9D6391F715 CRC64;

Query Match 28.6%; Score 54; DB 10; Length 315;
 Best local Similarity 42.3%; Pred. No. 9.6;
 Matches 11; Conservative 4; Mismatches 7; Indels 4; Gaps 1;

QY 8 YHLNVEQMWTTCQKPEFTNCVQPERE 33
 DB 24 HHLQV---TSCPDQATTNCTDQDRK 45

Search completed: March 18, 2004, 14:21:31
 Job time : 20.7738 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:15 ; Search time 44.4306 Seconds
(without alignments)
311.606 Million cell updates/sec

Title: US-09-941-314-13
272

Perfect score: 1 DQYKXSDDKYHFRIFRVLK.....EYHNVEMQWTCCKPPTN 49
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_29Jan04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	272	100.0	49 5 AAU79863	AAU79863 Human cys
2	272	100.0	115 5 AAU79853	AAU79853 Human cys
3	272	100.0	117 5 AAU79854	AAU79854 Human cys
4	272	100.0	137 5 AAU79852	AAU79852 Human cys
5	266	97.8	80 5 AAU79865	AAU79865 Human cys
6	254	93.4	46 5 AAU79860	AAU79860 Human cys
7	248	91.2	52 5 AAU79864	AAU79864 Human cys
8	154	56.6	59 5 AAU79866	AAU79866 Human cys
9	149	54.8	33 5 AAU79862	AAU79862 Human cys
10	143	52.6	27 5 AAU79859	AAU79859 Human cys
11	143	52.6	35 5 AAU79858	AAU79858 Human cys
12	138	50.7	24 5 AAU79861	AAU79861 Human cys
13	128	47.1	50 4 AAM15096	AAM15096 Peptide #
14	128	47.1	50 4 ABB34086	ABB34086 Peptide #
15	128	47.1	50 4 AAM27545	AAM27545 Peptide #
16	128	47.1	50 4 ABB33389	ABB33389 Peptide #
17	128	47.1	50 4 ABB28913	ABB28913 Peptide #
18	128	47.1	50 4 ABB19524	ABB19524 Protein #
19	128	47.1	50 4 AAM67252	AAM67252 Human bon
20	128	47.1	50 4 AAM54871	AAM54871 Human bra
21	128	47.1	50 4 ABB48915	ABB48915 Human liv
22	128	47.1	50 4 AAM02833	AAM02833 Peptide #
23	128	47.1	50 5 ABB36903	ABB36903 Human pep
24	127	46.7	92 2 AAM78259	AAM78259 Fragment
25	127	46.7	113 6 ADA57563	ADA57563 Human sec

ALIGNMENTS

26	127	46.7	113	6	ADA41457	ADA41457 Human sec
27	127	46.7	113	7	ADC74577	ADC74577 Human sec
28	127	46.7	113	7	ADD38088	ADD38088 Human sec
29	127	46.7	114	2	AAW78153	AAW78153 Human sec
30	127	46.7	123	2	AAW78260	AAW78260 Fragment
31	127	46.7	142	2	AAW78258	AAW78258 Human cys
32	127	46.7	142	4	AAE02405	AAE02405 Human cys
33	127	46.7	142	4	AAE04434	AAE04434 Human cys
34	127	46.7	142	6	ADA57231	ADA57231 Human sec
35	127	46.7	142	6	ADA41112	ADA41112 Human sec
36	127	46.7	142	7	ADC74335	ADC74335 Human sec
37	127	46.7	142	7	ADD37980	ADD37980 Human pro
38	127	46.7	142	7	ADD46706	ADD46706 Human pro
39	127	46.7	142	7	ADD46710	ADD46710 Human pro
40	122	44.9	141	3	AAV95576	AAV95576 Murine cy
41	122	44.9	141	4	AAE02403	AAE02403 Murine cy
42	122	44.9	141	4	AAE04432	AAE04432 Mouse tes
43	120	44.1	142	4	AAE02404	AAE02404 Murine cy
44	120	44.1	142	4	AAE04433	AAE04433 Mouse cys
45	120	44.1	143	6	ADA14374	ADA14374 Mouse spe

RESULT 1

AAU79863 standard; peptide; 49 AA.

AAU79863;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcy8) antigenic fragment #11.

Cystatin-8; Zcy8; cancer; procoagulant protein; thrombosis;

spermatogenesis; seminal fluid viscosity; cryopreserved sperm;

sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026668.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting spermatogenesis, and inhibiting cancer procoagulant protein which leads to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97-98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcy8) polypeptide (I). (I) is useful for: inhibiting cancer procoagulant protein in an individual and thus inhibiting the thrombotic events associated with cancer; promoting spermatogenesis, modulating seminal fluid viscosity, enhancing viability of cryopreserved sperm, sperm motility and fertilisation; and as antigenic peptides to generate antibodies. Zcy8 is useful as research reagent for characterising sites of interaction between Zcy8 and its receptor. Zcy8 is useful in enhancing fertilisation during assisted reproduction in humans and in animals. Anti-(I) antibodies are useful to screen biological samples like blood, urine, saliva, tissue biopsy and autopsy material in vitro for the presence of Zcy8. The antibodies are also useful to isolate large

CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
 XX
 SQ Sequence 49 AA;

Query Match 100.0%; Score 272; DB 5; Length 49;
 Best Local Similarity 100.0%; Pred. No. 4e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DOWNKESDDKXHFRIFFVLKVGQROVTDHLEHNLVEMQMTTCKQPEPTN 49
 Db 1 DOWNKESDDKXHFRIFFVLKVGQROVTDHLEHNLVEMQMTTCKQPEPTN 49

RESULT 2

AAU79853
 ID AAU79853 standard; protein; 115 AA.

AC AAU79853;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #1.

KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

PN WO200220567-A2.

PD 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026686.

PR 01-SEP-2000; 2000US-0230230P.

PA (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis; modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The

CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and
 CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
 CC This sequence represents an antigenic fragment of human cystatin-8
 CC (Zcys8)
 XX
 SQ Sequence 115 AA;

Query Match 100.0%; Score 272; DB 5; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1.1e-29;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DOWNKESDDKXHFRIFFVLKVGQROVTDHLEHNLVEMQMTTCKQPEPTN 49
 Db 30 DOWNKESDDKXHFRIFFVLKVGQROVTDHLEHNLVEMQMTTCKQPEPTN 78

RESULT 3

AAU79854
 ID AAU79854 standard; protein; 117 AA.

AC AAU79854;

DT 15-JUL-2002 (first entry)

DE Human cystatin-8 (Zcys8) antigenic fragment #2.

KM Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
 KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
 KM sperm motility; fertilisation; antigenic fragment.

OS Homo sapiens.

PN WO200220567-A2.

PD 14-MAR-2002.

PF 29-AUG-2001; 2001WO-US026686.

PR 01-SEP-2000; 2000US-0230230P.

PA (ZYMO) ZYMOGENETICS INC.

PI Holloway JL, Gao Z, Bishop PD;

DR WPI; 2002-383044/41.

PT Novel isolated mammalian cystatin-8 polypeptide useful for promoting
 PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
 PT to inhibition of thrombotic events associated with cancer.

PS Claim 2; Page 94-95; 100pp; English.

XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
 CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
 CC protein in an individual and thus inhibiting the thrombotic events
 CC associated with cancer; promoting spermatogenesis; modulating seminal
 CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
 CC motility and fertilisation; and as antigenic peptides to generate
 CC antibodies. Zcys8 is useful as research reagent for characterising sites
 CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
 CC enhancing fertilisation during assisted reproduction in humans and in
 CC animals. Anti-(I) antibodies are useful to screen biological samples like
 CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
 CC presence of Zcys8. The antibodies are also useful to isolate large
 CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
 CC The polynucleotide encoding (I) is useful to detect and to localise the
 CC expression of a Zcys8 gene in a biological sample and Zcys8
 CC oligonucleotide probes are useful for in vivo diagnosis. The
 CC polynucleotide encoding (I) is useful in determining whether a subject's
 CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
 CC copy number changes, insertions, deletions, restriction site changes and

CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic fragment of human cystatin-8
CC (Zcys8)
XX Sequence 117 AA;
SQ
Query Match 100.0%; Score 272; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKXHPRIFRVLKVRQVTDHLEHYLVNEMQWTTCKRPETTN 49
32 DQYNKESDDKXHPRIFRVLKVRQVTDHLEHYLVNEMQWTTCKRPETTN 80
Db
RESULT 4
AAU79852
ID AAU79852 standard; protein; 137 AA.
XX
AC AAU79852;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8).
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KM sperm motility; fertilisation.
XX
OS Homo sapiens.
XX
PN MO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001MO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
PI
PI Holloway JL, Gao Z, Bishop PD;
XX
XX WPI; 2002-383044/41.
DR N-PSDB; ABK49522.
XX
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 93-94; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This is the amino acid sequence of human cystatin-8 (Zcys8)

XX
SQ Sequence 137 AA;
SQ
Query Match 100.0%; Score 272; DB 5; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.4e-29;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKXHPRIFRVLKVRQVTDHLEHYLVNEMQWTTCKRPETTN 49
52 DQYNKESDDKXHPRIFRVLKVRQVTDHLEHYLVNEMQWTTCKRPETTN 100
Db
RESULT 5
AAU79865
ID AAU79865 standard; peptide; 80 AA.
XX
AC AAU79865;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cystatin-8 (Zcys8) antigenic fragment #13.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KM sperm motility; fertilisation; antigenic peptide.
XX
OS Homo sapiens.
XX
PN MO200220567-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001MO-US026868.
XX
PR 01-SEP-2000; 2000US-0230230P.
XX
PA (ZYMO) ZYMOGENETICS INC.
PI
PI Holloway JL, Gao Z, Bishop PD;
XX
XX WPI; 2002-383044/41.
DR
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
PS Claim 2; Page 98; 100pp; English.
XX
CC The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
SQ Sequence 80 AA;
SQ
Query Match 97.8%; Score 266; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 4.9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QVNESDDKXHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 49
DB 1 QVNESDDKXHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 48

RESULT 6

AAU79860
ID AAU79860 standard; peptide; 46 AA.

AAU79860;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #8.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
spermatogenesis, and inhibiting cancer procoagulant protein which leads
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 97; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
protein in an individual and thus inhibiting the thrombotic events
associated with cancer; promoting spermatogenesis, modulating seminal
fluid viscosity, enhancing viability of cryopreserved sperm, sperm
motility and fertilisation; and as antigenic peptides to generate
antibodies. Zcys8 is useful as research reagent for characterising sites
of interaction between Zcys8 and its receptor. Zcys8 is useful in
enhancing fertilisation during assisted reproduction in humans and in
animals. Anti-(I) antibodies are useful to screen biological samples like
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
presence of Zcys8. The antibodies are also useful to isolate large
quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
The polynucleotide encoding (I) is useful to detect and to localise the
expression of a Zcys8 gene in a biological sample and Zcys8
oligonucleotide probes are useful for in vivo diagnosis. The
oligonucleotide encoding (I) is useful in determining whether a subject's
chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
copy number changes, insertions, deletions, restriction site changes and
rearrangements and genetic alterations that inactivate the Zcys8 gene.
This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 46 AA;

Query Match 93.4%; Score 254; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NKESDDKXHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 49

DB 1 NKESDDKXHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 46

RESULT 7

AAU79864
ID AAU79864 standard; peptide; 52 AA.

AAU79864;

15-JUL-2002 (first entry)

Human cystatin-8 (Zcys8) antigenic fragment #12.

Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
sperm motility; fertilisation; antigenic peptide.

Homo sapiens.

MO200220567-A2.

14-MAR-2002.

29-AUG-2001; 2001WO-US026868.

01-SEP-2000; 2000US-0230230P.

(ZYMO) ZYMOGENETICS INC.

Holloway JL, Gao Z, Bishop PD;

WPI; 2002-383044/41.

Novel isolated mammalian cystatin-8 polypeptide useful for promoting
spermatogenesis, and inhibiting cancer procoagulant protein which leads
to inhibition of thrombotic events associated with cancer.

Claim 2; Page 98; 100pp; English.

The invention describes an isolated mammalian cystatin-8 (Zcys8)
polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
protein in an individual and thus inhibiting the thrombotic events
associated with cancer; promoting spermatogenesis, modulating seminal
fluid viscosity, enhancing viability of cryopreserved sperm, sperm
motility and fertilisation; and as antigenic peptides to generate
antibodies. Zcys8 is useful as research reagent for characterising sites
of interaction between Zcys8 and its receptor. Zcys8 is useful in
enhancing fertilisation during assisted reproduction in humans and in
animals. Anti-(I) antibodies are useful to screen biological samples like
blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
presence of Zcys8. The antibodies are also useful to isolate large
quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
The polynucleotide encoding (I) is useful to detect and to localise the
expression of a Zcys8 gene in a biological sample and Zcys8
oligonucleotide probes are useful for in vivo diagnosis. The
oligonucleotide encoding (I) is useful in determining whether a subject's
chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
copy number changes, insertions, deletions, restriction site changes and
rearrangements and genetic alterations that inactivate the Zcys8 gene.
This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)

Sequence 52 AA;

Query Match 91.2%; Score 248; DB 5; Length 52;
Best Local Similarity 100.0%; Pred. No. 9.2e-27;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KESDDKXHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 49
DB 1 KESDDKXHFRIFRVLKQROVTDHLEVHLNVEMQWTTCKPEPTN 45

XX	AAU79866	standard; peptide; 59 AA.
XX	AAU79866;	
XX	15-JUL-2002	(first entry)
XX	Human cystatin-8 (Zcys8) antigenic fragment #14.	
XX	Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;	
XX	seminal fluid viscosity; cryopreserved sperm;	
XX	sperm motility; fertilisation; antigenic peptide.	
XX	Homo sapiens.	
XX	WO200220567-A2.	
XX	14-MAR-2002.	
XX	29-AUG-2001; 2001WO-US026868.	
XX	01-SEP-2000; 2000US-0230230P.	
XX	(ZYMO) ZYMOGENETICS INC.	
XX	Holloway JL, Gao Z, Bishop PD;	
XX	WPI; 2002-383044/41.	
XX	Novel isolated mammalian cystatin-8 polypeptide useful for promoting	
XX	spermatogenesis, and inhibiting cancer procoagulant protein which leads	
XX	to inhibition of thrombotic events associated with cancer.	
XX	Claim 2; Page 99; 100pp; English.	
XX	The invention describes an isolated mammalian cystatin-8 (Zcys8)	
XX	polypeptide (I). (I) is useful for: inhibiting cancer procoagulant	
XX	protein in an individual and thus inhibiting the thrombotic events	
XX	associated with cancer; promoting spermatogenesis, modulating seminal	
XX	fluid viscosity, enhancing viability of cryopreserved sperm, sperm	
XX	motility and fertilisation; and as antigenic peptides to generate	
XX	antibodies. Zcys8 is useful as research reagent for characterising sites	
XX	of interaction between Zcys8 and its receptor. Zcys8 is useful in	
XX	enhancing fertilisation during assisted reproduction in humans and in	
XX	animals. Anti-(I) antibodies are useful to screen biological samples like	
XX	blood, urine, saliva, tissue biopsy and autopsy material in vitro for the	
XX	presence of Zcys8. The antibodies are also useful to isolate large	
XX	quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.	
XX	The polynucleotide encoding (I) is useful to detect and to localise the	
XX	expression of a Zcys8 gene in a biological sample and Zcys8	
XX	oligonucleotide probes are useful for in vitro diagnosis. The	
XX	polynucleotide encoding (I) is useful in determining whether a subject's	
XX	chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene	
XX	copy number changes, insertions, deletions, restriction site changes and	
XX	rearrangements and genetic alterations that inactivate the Zcys8 gene.	
XX	This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)	
XX	Sequence 59 AA;	
XX	Query Match	56.6%; Score 154; DB 5; Length 59;
XX	Best Local Similarity	100.0%; Pred. No. 1.2e-13;
XX	Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	23 RQVTDHLEFHLNVEMQWTTCKPPTN	49
XX	1 RQVTDHLEFHLNVEMQWTTCKPPTN	27
XX	AAU79862 standard; peptide; 33 AA.	

AC	AAU9862;
XX	
DT	15-JUL-2002 (first entry)
DE	Human cystatin-8 (Zcys8) antigenic fragment #10.
KM	Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW	spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KX	sperm motility; fertilisation; antigenic peptide.
OS	Homo sapiens.
XX	
PN	MO200220567-A2.
PD	14-MAR-2002.
PF	29-AUG-2001; 2001WO-US026868.
PR	01-SEP-2000; 2000US-0230230P.
PA	(ZYMO) ZYMOGENETICS INC.
PI	Holloway JL, Gao Z, Bishop PD;
PJ	WPI; 2002-383044/41.
PT	Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PP	spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT	to inhibition of thrombotic events associated with cancer.
PS	Claim 2; Page 97; 100pp; English.
XX	
CC	The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC	polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC	protein in an individual and thus inhibiting the thrombotic events
CC	associated with cancer; promoting spermatogenesis, modulating seminal
CC	fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC	motility and fertilisation; and as antigenic peptides to generate
CC	antibodies. Zcys8 is useful as research reagent for characterising sites
CC	of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC	enhancing fertilisation during assisted reproduction in humans and in
CC	animals. Anti-(I) antibodies are useful to screen biological samples like
CC	blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC	presence of Zcys8. The antibodies are also useful to isolate large
CC	quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC	The polynucleotide encoding (I) is useful to detect and to localise the
CC	expression of a Zcys8 gene in a biological sample and Zcys8
CC	oligonucleotide probes are useful for in vivo diagnosis. The
CC	polynucleotide encoding (I) is useful in determining whether a subject's
CC	chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC	copy number changes, insertions, deletions, restriction site changes and
CC	rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC	This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX	
SQ	Sequence 33 AA;
Query Match	54.8%; Score 149; DB 5; Length 33;
Best Local Similarity	100.0%; Pred. NO. 2.9e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Qy	24 QVTDHLEYLHNVENMQTTCCQKPEFTTN 49
Db	1 QVTDHLEYLHNVENMQTTCCQKPEFTTN 26
RESULT 10	
ID	AAU79859 standard; peptide; 27 AA.
XX	
XC	AAU79859;
XX	
DT	15-JUL-2002 (first entry)
XX	

```

DE Human cystatin-8 (Zcys8) antigenic fragment #7
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW sperm motility; fertilisation; antigenic peptide.
XX
XX Homo sapiens.
OS
XX
XX WO200220567-A2.
FN
XX
XX 14-MAR-2002.
PD
XX
XX 29-UG-2001; 2001WO-US026868.
PF
XX
XX 01-SEP-2000; 2000US-0230230P.
PR
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA
XX
XX Holloway JL, Gao Z, Bishop PD;
PI
XX
XX WPI; 2002-383044/41.
DR
XX
XX
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
XX
XX Claim 2; Page 96; 100pp; English.
PS
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(I) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (I) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (I) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
XX Sequence 27 AA;
SQ
XX
XX Query Match 52.6%; Score 143; DB 5; Length 27;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-12;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX 1 DQYNKESDDKYHFRIFRYLKVQRYQYTD 27
QY |||||
XX |||||
DQ 1 DQYNKESDDKYHFRIFRYLKVQRYQYTD 27
XX
XX RESULT 11
XX ID AAU79858
XX AC AAU79858;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human cystatin-8 (Zcys8) antigenic fragment #6.
XX
XX Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KM spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
KW

```

XX	sperm motility; fertilisation; antigenic peptide.
OS	Homo sapiens.
XX	
PN	WO200220567-A2.
XX	
PD	14-MAR-2002.
XX	
PF	29-AUG-2001; 2001WO-US026668.
XX	
PR	01-SEP-2000; 2000US-0230230P.
XX	
PA	(Zymo) ZYMOGENETICS INC.
XX	
P1	Holloway JL, Gao Z, Bishop PD;
DR	WPI; 2002-383044/41.
XX	
PT	Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT	spermatogenesis, and inhibiting cancer procoagulant protein which leads
XX	to inhibition of chromocytic events associated with cancer.
XX	
PS	Claim 2; Page 96; 100p; English.
XX	
CC	The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC	polypeptide (I). (I) is useful for: inhibiting cancer procoagulant
CC	protein in an individual and thus inhibiting the thrombotic events
CC	associated with cancer; promoting spermatogenesis, modulating seminal
CC	fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC	motility and fertilisation; and as antigenic peptides to generate
CC	antibodies. Zcys8 is useful as research reagent for characterising sites
CC	of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC	enhancing fertilisation during assisted reproduction in humans and in
CC	animals. Anti-(I) antibodies are useful to screen biological samples like
CC	blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC	presence of Zcys8. The antibodies are also useful to isolate large
CC	quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC	The polynucleotide encoding (I) is useful to detect and to localise the
CC	expression of a Zcys8 gene in a biological sample and Zcys8
CC	oligonucleotide probes are useful for in vivo diagnosis. The
CC	polynucleotide encoding (I) is useful in determining whether a subject's
CC	chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC	copy number changes, insertions, deletions, restriction site changes and
CC	rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC	This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX	
SO	Sequence 35 AA;
XX	
Query Match	52.6%; Score 143; DB 5; Length 35;
Best Local Similarity	100.0%; Pred. No. 2.le-12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 DQYKESDQKHYFRIFRYLAKYQROQVTD 27
DB	9 DQYKESDQKHYFRIFRYLAKYQROQVTD 35
XX	
RESULT 12	
ID	AAU79861
XX	AAU79861 standard; peptide; 24 AA.
AC	AAU79861;
XX	
DT	15-JUL-2002 (first entry)
XX	
DE	Human cystatin-8 (Zcys8) antigenic fragment #9.
XX	
XX	Cystatin-8; Zcys8; cancer; procoagulant protein; thrombosis;
KW	spermatogenesis; seminal fluid viscosity; cryopreserved sperm;
XX	sperm motility; fertilisation; antigenic peptide.
XX	
OS	Homo sapiens.
XX	

```

PN WO200220567-A2.
XX
XX 14-MAR-2002.
PD
XX
XX 29-AUG-2001; 2001WO-US026868.
PF
XX
XX 01-SEP-2000; 2000US-0230230P.
PR
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA
XX
XX Holloway JL, Gao Z, Bishop PD;
PI
XX
XX MPI, 2002-383044/41.
DR
XX
XX
XX
XX Novel isolated mammalian cystatin-8 polypeptide useful for promoting
PT spermatogenesis, and inhibiting cancer procoagulant protein which leads
PT to inhibition of thrombotic events associated with cancer.
XX
XX
PS Claim 2; Page 97; 100pp; English.
XX
XX The invention describes an isolated mammalian cystatin-8 (Zcys8)
CC polypeptide (1). (1) is useful for: inhibiting cancer procoagulant
CC protein in an individual and thus inhibiting the thrombotic events
CC associated with cancer; promoting spermatogenesis, modulating seminal
CC fluid viscosity, enhancing viability of cryopreserved sperm, sperm
CC motility and fertilisation; and as antigenic peptides to generate
CC antibodies. Zcys8 is useful as research reagent for characterising sites
CC of interaction between Zcys8 and its receptor. Zcys8 is useful in
CC enhancing fertilisation during assisted reproduction in humans and in
CC animals. Anti-(1) antibodies are useful to screen biological samples like
CC blood, urine, saliva, tissue biopsy and autopsy material in vitro for the
CC presence of Zcys8. The antibodies are also useful to isolate large
CC quantities of Zcys8 protein and DNA sequences that encode Zcys8 genes.
CC The polynucleotide encoding (1) is useful to detect and to localise the
CC expression of a Zcys8 gene in a biological sample and Zcys8
CC oligonucleotide probes are useful for in vivo diagnosis. The
CC polynucleotide encoding (1) is useful in determining whether a subject's
CC chromosomes contain a mutation in the Zcys8 gene like aneuploidy, gene
CC copy number changes, insertions, deletions, restriction site changes and
CC rearrangements and genetic alterations that inactivate the Zcys8 gene.
CC This sequence represents an antigenic peptide of human cystatin-8 (Zcys8)
XX
XX
SO Sequence 24 AA;
XX
XX
XX Query Match 50.7%; Score 138; DB 5; Length 24;
XX Best Local Similarity 100.0%; Pred. No. 6.7e-12;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX 23 RQVTDHLEYHLNVEQMWTTCQKPE 46
XX |||||
XX 1 RQVTDHLEYHLNVEQMWTTCQKPE 24
XX
XX
XX RESULT 13
XX AAM15096
XX ID AAM15096 standard; protein; 50 AA.
XX
XX AAM15096;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #1530 encoded by probe for measuring cervical gene expression.
XX
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX OS Homo sapiens.
XX
XX WO200157278-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US000670.
XX
XX

```

XX	04-FEB-2000; 2000US-0180312P.	
PR	26-MAY-2000; 2000US-0207456P.	
PR	30-JUN-2000; 2000US-00608408.	
PR	03-AUG-2000; 2000US-00632366.	
PR	21-SEP-2000; 2000US-0234687P.	
PR	27-SEP-2000; 2000US-0236359P.	
PR	04-OCT-2000; 2000GB-00024263.	
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-488901/53.	
XX		
PT	Human genome-derived single exon nucleic acid probes useful for analyzing	
XX	gene expression in human cervical epithelial cells.	
PS	Claim 27; SEQ ID NO 19922; 487bp; English.	
XX		
CC	The present invention relates to human single exon nucleic acid probes	
CC	(SNP: see AA110068-AA128459). The present sequence is a peptide encoded	
CC	by one such probe. The SNPs are derived from human HeLa cells. The SNPs	
CC	can be used to produce a single exon microarray, which can be used for	
CC	measuring human gene expression in a sample derived from human cervical	
CC	epithelial cells. By measuring gene expression, the probes are therefore	
CC	useful in grading and/or staging of diseases of the cervix, notably	
CC	cervical cancer. Note: The sequence data for this patent did not form	
CC	part of the printed specification, but was obtained in electronic format	
CC	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
SQ	Sequence 50 AA;	
	Query Match 47.1%; Score 128; DB 4; Length 50;	
	Best Local Similarity 100.0%; Pred.No.3,9e-10;	
	Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 DQYNKESDDKYHFRIFPVLYKVRQ 24	
DB	27 DQYNKESDDKYHFRIFPVLYKVRQ 50	
	RESULT 14	
	ABB34086	
ID	ABB34086 standard; peptide; 50 AA.	
XX		
AC	ABB34086;	
XX		
DT	04-FEB-2002 (first entry)	
XX		
DE	Peptide #1592 encoded by human foetal liver single exon probe.	
XX		
XX	Human; foetal liver; gene expression; single exon nucleic acid probe.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157277-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	30-JAN-2001; 2001WO-US000669.	
XX		
PR	04-FEB-2000; 2000US-0180312P.	
PR	26-MAY-2000; 2000US-0207456P.	
PR	30-JUN-2000; 2000US-00608408.	
PR	03-AUG-2000; 2000US-00632366.	
PR	21-SEP-2000; 2000US-0234687P.	
PR	27-SEP-2000; 2000US-0236359P.	
PR	04-OCT-2000; 2000GB-00024263.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	

```

XX WPI; 2001-483447/52.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 26721; 639pp + Sequence Listing; English.
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human fetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;
Query Match 47.1%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKXHFRIFRVLKVRQ 24
Db 27 DQYNKESDDKXHFRIFRVLKVRQ 50
RESULT 15
AAM27545
ID AAM27545 standard; protein; 50 AA.
XX
AC AAM27545;
XX
DT 17-OCT-2001 (first entry)
DE Peptide #1582 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR,
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 27814; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see A131315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of

```

```

CC human genetic disorders
XX
SQ Sequence 50 AA;
Query Match 47.1%; Score 128; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DQYNKESDDKXHFRIFRVLKVRQ 24
Db 27 DQYNKESDDKXHFRIFRVLKVRQ 50
Search completed: March 18, 2004, 14:15:00
Job time : 44.4306 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:07:11, Search time 12.5417 Seconds
(without alignments)
201.701 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272
Sequence: 1 DQNKESDCKYHFRIVLK.....EHLNEMQWTCQKPEPTN 49

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/2/1aa/5A COMB pep.*
2: /cgn2_6/ptodata/2/1aa/6A COMB pep.*
3: /cgn2_6/ptodata/2/1aa/6B COMB pep.*
4: /cgn2_6/ptodata/2/1aa/6C COMB pep.*
5: /cgn2_6/ptodata/2/1aa/6D COMB pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	46.7	142	3	US-09-431-480-4
2	127	46.7	142	3	US-09-617-302-4
3	122	44.9	141	3	US-09-431-480-2
4	122	44.9	141	3	US-09-617-302-2
5	120	44.1	142	3	US-09-431-480-3
6	120	44.1	142	3	US-09-617-302-3
7	112	41.2	127	4	US-08-849-303-19
8	108	39.7	140	4	US-09-886-319A-46
9	108	39.7	140	4	US-09-886-319A-48
10	104	38.2	140	3	US-09-431-480-5
11	104	38.2	140	3	US-09-617-302-5
12	104	38.2	140	4	US-08-849-303-18
13	103	37.9	112	4	US-08-849-303-16
14	103	37.9	116	4	US-09-775-932-16
15	103	37.9	118	4	US-09-775-932-24
16	103	37.9	120	4	US-09-775-932-2
17	103	37.9	120	6	US-08-791-532-4
18	103	37.9	139	2	US-08-791-532-4
19	103	37.9	139	4	US-09-314-777-4
20	103	37.9	139	4	US-08-849-303-15
21	103	37.9	145	2	US-08-832-535-11
22	103	37.9	146	2	US-08-791-532-3
23	103	37.9	146	2	US-08-744-138-3
24	103	37.9	146	3	US-09-019-485-4
25	103	37.9	146	3	US-09-314-777-3
26	103	37.9	146	3	US-09-431-480-6
27	103	37.9	146	3	US-09-617-302-6

28	103	37.9	146	4	US-09-241-376-3	Sequence 3, Appl
29	103	37.9	146	4	US-09-528-436B-3	Sequence 3, Appl
30	103	37.9	146	4	US-09-886-319A-47	Sequence 47, Appl
31	103	37.9	146	4	US-09-940-497-3	Sequence 3, Appl
32	103	37.9	146	4	US-09-976-594-37	Sequence 37, Appl
33	103	37.9	146	4	US-08-849-303-17	Sequence 17, Appl
34	103	37.9	146	5	PCT-US95-07135-9	Sequence 9, Appl
35	103	37.9	146	6	5432264-6	Patent No. 5432264
36	93	34.2	121	4	US-09-775-932-8	Sequence 8, Appl
37	93	34.2	121	3	US-08-744-138-6	Sequence 6, Appl
38	93	34.2	141	4	US-09-241-376-6	Sequence 6, Appl
39	93	34.2	141	4	US-09-940-497-6	Sequence 6, Appl
40	93	34.2	141	4	US-08-849-303-24	Sequence 24, Appl
41	90	33.1	111	4	US-08-849-303-26	Sequence 26, Appl
42	89	32.7	111	4	US-09-775-932-18	Sequence 18, Appl
43	85	31.2	121	4	US-09-775-932-14	Sequence 14, Appl
44	85	31.2	128	4	US-09-775-932-12	Sequence 12, Appl
45	85	31.2	149	2	US-08-461-030C-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-09-431-480-4
; Sequence 4, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; EARLIER FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-4

Query Match      46.7%; Score 127; DB 3; Length 142;
Best Local Similarity 50.0%; Pred. No. 4.2e-11;
Matches 24; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

Cy      2 QYKESDCKYHFRIVLKQVOTDHLNEMQWTCQKPEPTN 49
Db      55 EYKESDCKYVFLVVKTLQALQVNTLLEVLINVEIARSDCKRPLSTN 102

RESULT 2
US-09-617-302-4
; Sequence 4, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 C1
; CURRENT APPLICATION NUMBER: US/09/617,302
; EARLIER FILING DATE: 2000-07-17
; EARLIER APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 142
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-617-302-4

```

Query Match	46.7%;	Score 127;	DB 3;	Length 142;
Best Local Similarity	50.0%;	Pred. No. 4,2e-11;		
Matches 24;	Conservative 12;	Mismatches 12;	Indels 0;	Gaps 0;

```
QY      2 QNKSDDKXHHRIERVLKNRQVTDHLEHYHNEMQMWTQQKEPTTN 49
         :|||||:| | : : | |||:| | :||| : :| |:| :||
Db      55 EYNKSEDKYVFLVKTLQAQLQVTNLLEYLDIVEIARSDCKRP LSTN 102
```

```

: RESULT 3
: US-09-431-480-2
: Sequence 2, Application US/09431480
: Patent No. 6235708
: GENERAL INFORMATION:
: APPLICANT: Holloway, James L.
: APPLICANT: Felchhaus, Andrew
: TITLE OF INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
: FILE REFERENCE: 98-72
: CURRENT APPLICATION NUMBER: US/09/431,480
: CURRENT FILING DATE: 1999-11-01
: EARLIER APPLICATION NUMBER: 60/109,217
: EARLIER FILING DATE: 1998-11-20
: EARLIER APPLICATION NUMBER: 60/156,382
: EARLIER FILING DATE: 1999-09-28
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 141
: TYPE: PRY
: ORGANISM: Homo sapiens
: US-09-431-480-2

```

Query Match	44.9%;	Score 122;	DB 3;	Length 141;
Best Local Similarity	51.2%;	Pred. No. 2.3e-10;		
Matches 22;	Conservative	9;	Mismatches 12;	Indels 0;
			Gaps	0;

```

QY      2 QYNKESDDKHYFRIRFLVKQGRQVTDHLLEYHLNVEMQNTTCOK 44
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     54 EYNKASNDLYNFRVVDILKSQEQITDSLEYLEYLVNIARTMCKK 96

```

```

RESULT 4
US-09-617-302-2
Sequence 2, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN- LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
US-09-617-302-2

```

	Query Match	44.9%;	Score 1.22;	DB 3;	Length 141;
	Best Local Similarity	51.2%;	Pred. No. 12,36-10;		
Matches	22;	Conservative	9;	Mismatches	12; Indels 0; Gaps 0;
Oy	2 QYNKESDDKHFRLEFVLAKOROVTDLEHLNLMEMQMTTCOK	44			
		:			
Db	54 EYKAASNDLNFRRVVDILIKSGEQITDLSFYLLFNINARTMCK	96			

RESULT 5
US-09-431-480-3
Sequence 3 Application IIS/09431480

```

1  TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN
2  FILE REFERENCE: 98-72
3  CURRENT APPLICATION NUMBER: US/09/431,480
4  CURRENT FILING DATE: 1999-11-01
5  EARLIER APPLICATION NUMBER: 60/109,217
6  EARLIER FILING DATE: 1998-11-20
7  EARLIER APPLICATION NUMBER: 60/156,382
8  EARLIER FILING DATE: 1999-09-28
9  NUMBER OF SEQ. ID NOS: 22
10 SOFTWARE: FastSeq for Windows Version 3.0
11 SEQ. ID NO. 3
12 LENGTH: 142
13 TYPE: PRT
14 ORGANISM: Mus musculus
15 US-09-431-480-3

```

Query Match	44.1%;	Score 120;	DB 3;	Length 142;
Best Local Similarity	43.2%;	Pred. No. 4.5e-10;		
Matches 19;	Conservative 14;	Mismatches 11;	Indels 0;	Gaps 0;

```
QY      2 QYNKESDDKYHFRIFRVLKVQRQVLDHLLEYHLNVEMQWTTCCRP 45
          :|||:||| |::| :||| :||| :||| :|||
DB     55 EYNKESEDKVFLVDKIILAKLTIDRMEYQIDVQISRSNCKP 98
```

RESULT 6
US-09-617-302-3
; Sequence 3, Application US/09617302

```

1  APPLICANT: Holloway, James L.
2  APPLICANT: Feldhaus, Andrew
3  TITLE OR INVENTION: TESTS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
4  FILE REFERENCE: 98-72 C1
5  CURRENT APPLICATION NUMBER: US/09/617,302
6  CURRENT FILING DATE: 2000-07-17
7  PRIOR APPLICATION NUMBER: 09/431,480
8  PRIOR FILING DATE: 1999-11-01
9  PRIOR APPLICATION NUMBER: 60/109,217
10 PRIOR FILING DATE: 1998-11-20
11 PRIOR APPLICATION NUMBER: 60/156,382
12 PRIOR FILING DATE: 1999-09-28
13 NUMBER OF SEQ ID NOS: 22
14 SOFTWARE: FastSeq for Windows Version 3.0
15 SEQ ID NO 3
16 LENGTH: 142
17 TYPE: PRT
18 ORGANISM: Mus musculus
19 US-09-617-302-3

```

Query Match	44.1%	Score 120;	DB 3;	Length 142;
Best Local Similarity	43.2%	Pred. No. 4,5e-10;		
Matches 19;	Conservative 14;	Mismatches 11;	Indels 0;	Gaps 0;
QY	2 QYNESDDKHFRLEFVLKQVRYVDHLEHLYNTNEMQMTQCKP	45		
DB	55 EYNSESDKRYLVLDKILHAKLTITDRNEVQIDVQISSNCKP	98		

RESULT 7

US-08-849-303-19

Sequence 19, Application US/08849303

Patent No. 6680424

GENERAL INFORMATION:

APPLICANT: Atkinson, Howard J.

APPLICANT: McPherson, Michael J.

APPLICANT: Urwin, Peter E.

TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS

NUMBER OF SEQUENCES: 79

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue, 4th Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/849,303

FILING DATE: 21-May-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1321-1-003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 127 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

US-08-849-303-19

PRIOR FILING DATE: 2000-08-01
 PRIOR APPLICATION NUMBER: DE 10030149.5
 PRIOR FILING DATE: 2000-06-20
 NUMBER OF SEQ ID NOS: 84
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 46
 LENGTH: 140
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

PRIOR FILING DATE: 2000-08-01
 PRIOR APPLICATION NUMBER: DE 10030149.5
 PRIOR FILING DATE: 2000-06-20
 NUMBER OF SEQ ID NOS: 84
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 46
 LENGTH: 140
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

PRIOR FILING DATE: 2000-08-01
 PRIOR APPLICATION NUMBER: DE 10030149.5
 PRIOR FILING DATE: 2000-06-20
 NUMBER OF SEQ ID NOS: 84
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 46
 LENGTH: 140
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-886-319A-46

US-09-8

NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 140
TYPE: PRT
ORGANISM: Mus musculus
US-09-431-480-5

Query Match 38.2%; Score 104; DB 3; Length 140;
Best Local Similarity 41.3%; Pred. No. 1e-07;
Matches 19; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKVQGVTDHLEHNLNEMQWTTCKPCT 47
DB 53 EYNGSNDAYHSRAIQVVRARQLVAGVNYFFDVEMGRTTCTKSQT 98

RESULT 11
US-09-617-302-5
Sequence 5, Application US/09617302
Patent No. 6245529
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Feldhaus, Andrew
TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
FILE REFERENCE: 98-72 C1
CURRENT APPLICATION NUMBER: US/09/617,302
CURRENT FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/431,480
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: 60/109,217
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/156,382
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 140
TYPE: PRT
ORGANISM: Mus musculus
US-09-617-302-5

Query Match 38.2%; Score 104; DB 3; Length 140;
Best Local Similarity 41.3%; Pred. No. 1e-07;
Matches 19; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKVQGVTDHLEHNLNEMQWTTCKPCT 47
DB 53 EYNGSNDAYHSRAIQVVRARQLVAGVNYFFDVEMGRTTCTKSQT 98

RESULT 12
US-08-849-303-18
Sequence 18, Application US/08849303
Patent No. 6680424
GENERAL INFORMATION:
APPLICANT: Atkinson, Howard J.
APPLICANT: McPherson, Michael J.
APPLICANT: Urwin, Peter E.
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-849-303-18

Query Match 38.2%; Score 104; DB 4; Length 140;
Best Local Similarity 41.3%; Pred. No. 1e-07;
Matches 19; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 2 QYNKESDXYHFRIFRYLKVQGVTDHLEHNLNEMQWTTCKPCT 47
DB 53 EYNGSNDAYHSRAIQVVRARQLVAGVNYFFDVEMGRTTCTKSQT 98

RESULT 13
US-08-849-303-16
Sequence 16, Application US/08849303
Patent No. 6680424
GENERAL INFORMATION:
APPLICANT: Atkinson, Howard J.
APPLICANT: McPherson, Michael J.
APPLICANT: Urwin, Peter E.
TITLE OF INVENTION: MODIFIED PROTEINASE INHIBITORS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,303
FILING DATE: 21-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1321-1-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

Search completed: March 18, 2004, 14:25:37

```
QY      2 QYNKESDDKYHFRIFRVLKQROVTDHLLEYHLNVEMQMWTTCQKE 46
        ::|||:||:::||:|||||:
Db     26 EFNKRSDAYQSRVVRVVRARKQVVGSMNYFLDVELGRITCTKSQ 70
```

? Patent No. 6534477
 ? GENERAL INFORMATION:
 ? APPLICANT: University of British Columbia
 ? TITLE OF INVENTION: Production and use of Modified Cystatins
 ? FILE REFERENCE: 58069
 ? CURRENT APPLICATION NUMBER: US/09/775,932
 ? CURRENT FILING DATE: 2001-02-02
 ? PRIOR APPLICATION NUMBER: CA99/00717
 ? PRIOR FILING DATE: 1999-08-05
 ? PRIOR APPLICATION NUMBER: 60/095,503
 ? PRIOR FILING DATE: 1998-08-05
 ? NUMBER OF SEQ ID NOS: 32
 ? SOFTWARE: PatentIn Ver. 2.0
 ? SEQ ID NO 16
 ? LENGTH: 116
 ? TYPE: prt
 ? ORGANISM: Gallus sp.
 ? US-09-775-932-16

Qy 2 QYNKESDDKYHFRIFRVLAKVQRÖVTDLHEYLHNVMÖWTTÇOK 44
:::|::|::|::|::|::|::|
Db 31 EYNRASNDKYSSRVVRVISAKRQLVSGIKYILOVEIGRTCPK 73

```

US-09-775-932-24
; Sequence 24, Application US/09775532
; Patent No. 6534477
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OR INVENTION: Production and use of Modified Cystatins
; FILE REFERENCE: 56069
; CURRENT APPLICATION NUMBER: US/09/775, 932
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 66/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Bos taurus
; US-09-775-932-24

```

```
OY      2 QYNKESDDKYHFRIFRYALQVÖVTDLHLHYLNVEMÖWTTCKQPE 46
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB     32 EFNKRSDAYQSRRVRVRVARQVSGMNYFLDVELGRTCTKSQ   76
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 18, 2004, 14:15:06 ; Search time 29.6528 Seconds

(without alignments)
427.913 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272

Sequence: 1 DQYNKESDDKYNHRIFRVLK.....EYHLNVMQWTTCKPETT 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep:*

13: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep:*

14: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep:*

15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep:*

17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	272	100.0	49	US-09-941-314-13	Sequence 13, Appl1
2	272	100.0	115	US-09-941-314-3	Sequence 3, Appl1
3	272	100.0	117	US-09-941-314-4	Sequence 4, Appl1
4	272	100.0	137	US-09-941-314-2	Sequence 2, Appl1
5	266	97.8	80	US-09-941-314-15	Sequence 15, Appl1
6	254	93.4	46	US-09-941-314-10	Sequence 10, Appl1
7	248	91.2	52	US-09-941-314-14	Sequence 14, Appl1
8	154	56.6	59	US-09-941-314-16	Sequence 16, Appl1
9	149	54.8	33	US-09-941-314-12	Sequence 12, Appl1
10	143	52.6	27	US-09-941-314-9	Sequence 9, Appl1
11	143	52.6	35	US-09-941-314-8	Sequence 8, Appl1
12	138	50.7	24	US-09-941-314-11	Sequence 11, Appl1
13	128	47.1	50	US-09-864-761-34822	Sequence 34822, A
14	128	47.1	50	US-09-864-761-48936	Sequence 48936, A
15	114	41.9	145	US-09-740-638-2	Sequence 2, Appl1

16	114	41.9	145	13	US-10-006-467-2	Sequence 2, Appl1
17	114	41.9	145	14	US-10-235-148-2	Sequence 2, Appl1
18	112	41.2	127	8	US-08-849-303-19	Sequence 19, Appl1
19	108	39.7	140	14	US-10-376-564-46	Sequence 46, Appl1
20	108	39.7	140	14	US-10-376-564-48	Sequence 48, Appl1
21	108	39.7	145	14	US-10-168-425-14	Sequence 14, Appl1
22	107	39.3	181	15	US-10-168-425-14	Sequence 14, Appl1
23	104	38.2	140	8	US-10-264-049-2608	Sequence 2608, Ap
24	103	37.9	112	8	US-08-849-303-18	Sequence 18, Appl1
25	103	37.9	112	8	US-08-849-303-16	Sequence 16, Appl1
26	103	37.9	116	9	US-09-775-932-16	Sequence 16, Appl1
27	103	37.9	118	9	US-09-775-932-24	Sequence 24, Appl1
28	103	37.9	120	9	US-09-775-932-2	Sequence 2, Appl1
29	103	37.9	139	8	US-08-849-303-15	Sequence 15, Appl1
30	103	37.9	139	8	US-09-969-834-4	Sequence 4, Appl1
31	103	37.9	146	8	US-08-849-303-17	Sequence 17, Appl1
32	103	37.9	146	9	US-09-940-497-3	Sequence 3, Appl1
33	103	37.9	146	9	US-09-969-834-3	Sequence 3, Appl1
34	103	37.9	146	14	US-10-329-428-3	Sequence 3, Appl1
35	103	37.9	146	14	US-10-376-564-47	Sequence 47, Appl1
36	94	34.6	165	9	US-09-740-638-5	Sequence 5, Appl1
37	94	34.6	165	13	US-10-006-467-5	Sequence 5, Appl1
38	93	34.2	48	9	US-10-235-148-5	Sequence 5, Appl1
39	93	34.2	121	9	US-09-941-314-17	Sequence 17, Appl1
40	93	34.2	141	8	US-09-775-932-8	Sequence 8, Appl1
41	93	34.2	141	8	US-08-849-303-24	Sequence 24, Appl1
42	90	33.1	141	8	US-09-940-497-6	Sequence 6, Appl1
43	89	32.7	111	8	US-08-849-303-26	Sequence 26, Appl1
44	85	31.2	121	9	US-09-775-932-18	Sequence 18, Appl1
45	85	31.2	128	9	US-09-775-932-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1

US-09-941-314-13

Sequence 13, Application US/09941314

Patent No. US20020142396A1

GENERAL INFORMATION:

APPLICANT: ZymoGenetics, Inc.

TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to

TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC

CURRENT APPLICATION NUMBER: US/09/941,314

CURRENT FILING DATE: 2001-08-29

PRIOR APPLICATION NUMBER: 60/230,230

PRIOR FILING DATE: 2001-09-01

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13

LENGTH: 49

TYPE: PRT

ORGANISM: Homo sapiens

US-09-941-314-13

Query Match 100.0%; Score 272; DB 9; Length 49;

Best Local Similarity 100.0%; Pred. No. 7.2e-29;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKYNHRIFRVLK...EYHLNVMQWTTCKPETT 49

Db 1 DQYNKESDDKYNHRIFRVLK...EYHLNVMQWTTCKPETT 49

RESULT 2

US-09-941-314-3

Sequence 3, Application US/09941314

Patent No. US20020142396A1

GENERAL INFORMATION:

APPLICANT: ZymoGenetics, Inc.

TITLE OF INVENTION: Mammalian Cystatin-B and Its Use to

TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein

FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-3

Query Match 100.0%; Score 272; DB 9; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.9e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49
DB 30 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 78

RESULT 3
US-09-941-314-4
Sequence 4, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-4

Query Match 100.0%; Score 272; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.9e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49
DB 32 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 80

RESULT 4
US-09-941-314-2
Sequence 2, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 137
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-2

Query Match 100.0%; Score 272; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.3e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49
DB 52 DQNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 100

RESULT 5
US-09-941-314-15
Sequence 15, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 80
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-15

Query Match 97.8%; Score 266; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.9e-28;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49
DB 1 QYNKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 48

RESULT 6
US-09-941-314-10
Sequence 10, Application US/09941314
Patent No. US20020142396A1
GENERAL INFORMATION:
APPLICANT: ZymoGenetics, Inc.
TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
TITLE OF INVENTION: Inhibit Cancer Procoagulant Protein
FILE REFERENCE: 00-81PC
CURRENT APPLICATION NUMBER: US/09/941.314
CURRENT FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: 60/230,230
PRIOR FILING DATE: 2001-09-01
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
US-09-941-314-10

Query Match 93.4%; Score 254; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.7e-26;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 49
DB 1 NKESDDKXHFRIFFVLKVRQVTDHLEYNLVEMQWTTCCQKPEPTN 46

RESULT 7
US-09-941-314-14
Sequence 14, Application US/09941314
Patent No. US20020142396A1

```

; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-14

Query Match      91.2%; Score 248; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.2e-25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 KESDDKYHFRIFRVLKVQROVTDHLEVHLNVEMQWTTCKPKETTN 49
Db      1 KESDDKYHFRIFRVLKVQROVTDHLEVHLNVEMQWTTCKPKETTN 45

RESULT 8
US-09-941-314-16
; Sequence 16, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-16

Query Match      56.6%; Score 154; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.7e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      23 QVTDHLEVHLNVEMQWTTCKPKETTN 49
Db      1 QVTDHLEVHLNVEMQWTTCKPKETTN 27

RESULT 9
US-09-941-314-12
; Sequence 12, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 33

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-12

Query Match      54.8%; Score 149; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 QVTDHLEVHLNVEMQWTTCKPKETTN 49
Db      1 QVTDHLEVHLNVEMQWTTCKPKETTN 26

RESULT 10
US-09-941-314-9
; Sequence 9, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-9

Query Match      52.6%; Score 143; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.7e-12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DQYKESDDKYHFRIFRVLKVQROVTD 27
Db      1 DQYKESDDKYHFRIFRVLKVQROVTD 27

RESULT 11
US-09-941-314-8
; Sequence 8, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-941-314-8

Query Match      52.6%; Score 143; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 7.6e-12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DQYKESDDKYHFRIFRVLKVQROVTD 27
Db      9 DQYKESDDKYHFRIFRVLKVQROVTD 35

```

```
RESULT 12
US-09-941-314-11
; Sequence 11, Application US/09941314
; Patent No. US20020142396A1
; GENERAL INFORMATION:
; APPLICANT: Zymogenetics, Inc.
; TITLE OF INVENTION: Mammalian Cystatin-8 and Its Use to
; FILE REFERENCE: 00-81PC
; CURRENT APPLICATION NUMBER: US/09/941,314
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/230,230
; PRIOR FILING DATE: 2001-09-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-314-11

Query Match          50.7%; Score 138; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.3e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      23 ROVTDLHLYHNVEMQWTTCKQPE 46
Db      1 ROVTDLHLYHNVEMQWTTCKQPE 24

RESULT 13
US-09-864-761-34822
; Sequence 34822, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34822
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109954.10
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: A120857.1, EVALUATE 5.00e-23
; OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUATE 1.00e-01
US-09-864-761-34822

Query Match          47.1%; Score 128; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DOYNKESDDKYHFRIFVLKYQRO 24
Db      27 DOYNKESDDKYHFRIFVLKYQRO 50

RESULT 14
US-09-864-761-48936
; Sequence 48936, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```


PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48936
LENGTH: 50
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL093677.18
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96
OTHER INFORMATION: EST HUMAN HIT: A1200857.1, EVALUE 5.00e-23
OTHER INFORMATION: SWISSPROT HIT: O60676, EVALUE 1.00e-01
US-09-864-761-48936

Query Match 47.1%; Score 128; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYNKESDDKTHFRIFRVKVRQ 24
DB 27 DQYNKESDDKTHFRIFRVKVRQ 50

RESULT 15
US-09-740-638-2
Sequence 2, Application US/09740638
Patent No. US20020006656A1
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
TITLE OF INVENTION: Zcyas: A Member of the Cystatin
TITLE OF INVENTION: Superfamily
FILE REFERENCE: 99-104
CURRENT APPLICATION NUMBER: US/09/740,638
CURRENT FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 145
TYPE: PRT
ORGANISM: Homo sapiens
US-09-740-638-2

Query Match 41.9%; Score 114; DB 9; Length 145;
Best Local Similarity 40.4%; Pred. No. 2.8e-07;
Matches 19; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

QY 3 YNKESDDKTHFRIFRVKVRQVTDHLEHIANEMOWTTCOKPETTN 49
DB 52 YNANSDTYLYRVORLIRSORLTTGVEYIVYKIGTKCKRNDTSN 98

Search completed: March 18, 2004, 14:30:54
Job time : 29.6528 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:06:15 ; Search time 9.43056 Seconds
(without alignments)
499.799 Million cell updates/sec

Title: US-09-941-314-13
Perfect score: 272
Sequence: 1 DQYKESDQKXFRIFRVLK.....EYHNVEMQWTCQKPEPTN 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	44.1	139	2 A45361	cystatin-related e
2	112	41.2	120	2 S10587	cystatin C - rat
3	112	41.2	127	2 S07085	cystatin C precurs
4	104	38.2	140	2 A36163	cystatin C precurs
5	103	37.9	112	1 UD80	cystatin - bovine
6	103	37.9	139	1 UDCH	cystatin precursor
7	103	37.9	146	1 UDHU	cystatin C precurs
8	93	34.2	141	2 B29632	cystatin SA precur
9	90	33.1	111	2 A28793	cystatin - puff ad
10	82	30.1	141	1 UDHUP2	cystatin SN precur
11	81	29.8	162	2 A43428	onchocystatin - ne
12	80	29.4	141	1 UDHUP1	cystatin S precurs
13	77	28.3	498	2 T31871	hypothetical prote
14	75	27.6	111	1 JC2040	cystatin - chum sa
15	75	27.6	132	1 JC4948	cystatin precursor
16	75	27.6	133	2 JC4536	cystatin precursor
17	74.5	27.4	142	2 A47142	cystatin D precurs
18	74.5	27.4	141	2 J01470	cystatin S precurs
19	64	23.5	436	1 KGB01	kininogen, LMW I p
20	64	23.5	621	1 KGB01	kininogen, LMW I p
21	63	23.2	438	2 A47702	glucan 1,3-beta-gl
22	63	23.2	438	2 T52149	beta-glucanase [im
23	62.5	23.0	448	2 JN0118	kininogen, LMW II
24	62	22.8	434	1 KGB01	kininogen, LMW II
25	62	22.8	619	1 KGB02	kininogen, LMW II
26	60	22.1	257	2 T03724	C-type cyclin - xi
27	59	21.7	139	2 B64005	hypothetical prote
28	59	21.7	325	2 F69784	conserved hypochet
29	59	21.7	427	1 KGHU1	kininogen, LMW pre

30	59	21.7	644	1 KGHU1	kininogen, LMW pre
31	58.5	21.5	246	2 H90539	ABC transporter at
32	58.5	21.5	382	2 G90540	hypothetical prote
33	58	21.3	47	2 S62326	cystatin C, paroti
34	57.5	21.1	428	2 A45568	highly immunoreact
35	57.5	21.1	1585	2 T19121	probable protein-t
36	57	21.0	288	2 T04401	glutathione-S-trans
37	56.5	20.8	483	2 S61975	secreted phosphop
38	56	20.6	200	2 I46051	recruited phosphop
39	56	20.6	588	2 G93552	L-fucose isomerase
40	56	20.6	588	2 A97117	L-fucose isomerase
41	56	20.6	617	2 S19254	nitrate reductase
42	55.5	20.4	164	2 T31026	hypothetical prote
43	55.5	20.4	527	2 E90740	probable enzyme [i
44	55.5	20.4	527	2 G85590	probable enzyme yb
45	55.5	20.4	527	2 G64818	probable membrane

ALIGNMENTS

RESULT 1
A45361
cystatin-related epididymal specific protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C/Accession: A45361
R/Cornwall, G.A.; Orgebin-Chriec, M.C.; Ham, S.R.
Mol. Endocrinol. 6, 1653-1664, 1992
A/Title: The CRIS gene, a unique testis-regulated gene related to the cystatin family is
A/Reference number: A45361; MUID:93078799; PMID:1280328
A/Accession: A45361
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-139 <COR>
A/Cross-references: GB:549926; NID:g260492; PIDN:AAC35390.1; PID:g260493
A/Note: sequence extracted from NCBI backbone (NCBI:118813)
C/Superfamily: cystatin; cystatin homology
F:28-139/Domain: cystatin homology <CYS>

Query Match 44.1%; Score 120; DB 2; Length 139;
Best Local Similarity 43.2%; Pred. No. 3.4e-08;
Matches 19; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 2 QYKESDQKXFRIFRVLKQVROVTHLEHNVEMQWTCQKPEPTN 45
Db 52 EYKESDQKXFRIFRVLKQVROVTHLEHNVEMQWTCQKPEPTN 95

RESULT 2
S10587
cystatin C - rat
C/Species: Rattus sp. (rat)
C/Date: 21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C/Accession: S10587
R/Barnard, F.; Barnard, A.; Faucher, D.; Capony, J.P.; Derancourt, J.; Billard, M.; Gauth
Biol. Chem. Hoppe-Seyler 371(Suppl.), 161-166, 1990
A/Title: Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation
A/Reference number: S10587; MUID:90380276; PMID:2400577
A/Accession: S10587
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-120 <ESN>
A/Note: 43-Asn was also found
A/Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 18-Ala
C/Superfamily: cystatin; cystatin homology
F:9-120/Domain: cystatin homology <CYS>

Query Match 41.2%; Score 112; DB 2; Length 120;
Best Local Similarity 43.5%; Pred. No. 3.1e-07;
Matches 20; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYKESDQKXFRIFRVLKQVROVTHLEHNVEMQWTCQKPEPTN 47

A:Reference number: S48159; MUID:95010016; PMID:7925354
 A:Accession: S48159
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 24-139 <NDS>
 R:Lieber, B.; Krieglstein, K.; Henechen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.
 FEBS Lett. 248, 162-168, 1999
 A:Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.
 A:Reference number: S04008; MUID:89252033; PMID:2721673
 A:Accession: S04008
 A:Molecule type: protein
 A:Residues: 97-114 <LAB>
 R:Colletta, R.; Bird, J.W.C.
 Gene 130, 175-181, 1993
 A:Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping
 A:Reference number: JN0789; MUID:9336172; PMID:8359684
 A:Accession: JN0789
 A:Molecule type: DNA
 A:Residues: 1-139 <CO2>
 A:Cross-references: GB:M95725
 A:Note: authors failed to translate the codon for residue 115-Tyr
 C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases
 C:GeneCis:
 A:Gene: Can
 A:Introns: 76/3; 114/3
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-139/Product: cystatin, long form #status experimental <CYLP>
 F:30-139/Domain: cystatin homology <CYS>
 F:32-139/Product: cystatin, short form #status experimental <CYS>
 F:76-80/Region: inhibitory #status predicted
 F:94-104,118-138/Diulfide bonds: #status experimental
 F:103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 37.9%; Score 103; DB 1; Length 139;
 Best Local Similarity 44.2%; Pred. No. 5.3e-06;
 Matches 19; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYNKESDKYHFRIFVLKVRQVTDHLEHLNEMQWTCOK 44
 Db 54 EYNRASNDKYSRVVRVIAKRLQVSGIKYILQVEIGRTCPK 96

RESULT 7
 UDHU
 Cystatin C precursor [validated] - human
 N:Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma
 C:Species: Homo sapiens (man)
 C:Date: 06-Jul-1982 #sequence revision 31-Mar-1991 #text change 08-Dec-2000
 A:Accession: S10216; S00004; JI0095; A33400; S02751; A01270; A25434; S12288; A32732; A60
 R:Abrahamson, M.; Olafsson, I.; Palstodt, A.; Ulvbaeck, M.; Lundvall, A.; Jansson, O.
 Biochem. J. 268, 287-294, 1990
 A:Title: Structure and expression of the human cystatin C gene.
 A:Reference number: S10216; MUID:90303202; PMID:2353674
 A:Accession: S10216
 A:Molecule type: DNA
 A:Residues: 1-146 <AB1>
 A:Cross-references: EMBL:X52255; NID:930257; PIDN:CAA36497.1; PID:g296643
 R:Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundvall, A.
 FEBS Lett. 216, 229-233, 1987
 A:Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the
 A:Reference number: S00004; MUID:87219149; PMID:3495457
 A:Accession: S00004
 A:Molecule type: mRNA
 A:Residues: 1-146 <AB2>
 R:Levy, E.; Lopez-Otin, C.; Ghiso, J.; Gellner, D.; Frangione, B.
 J. Exp. Med. 169, 1771-1778, 1989
 A:Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a
 A:Reference number: JI0095; MUID:89235594; PMID:2541223
 A:Accession: JI0095
 A:Molecule type: DNA

A:Residues: 1-146 <LEV>
 A:Cross-references: GB:X61681; NID:930367; PIDN:CAA43956.2; PID:g4490944
 A:Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit
 e)
 R:Salton, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada, H.
 Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989
 A:Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which
 A:Reference number: A33400; MUID:89350949; PMID:2764935
 A:Accession: A33400
 A:Molecule type: DNA
 A:Residues: 1-24, 'T', 26-146 <SA1>
 A:Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:9181385; PIDN:AAA52164.1; PID:g
 R:Ghisso, J.; Cowan, N.; Frangione, B.
 Biol. Chem. Hoppe-Seyler 369, 205-209, 1988
 A:Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron
 A:Reference number: S02751; MUID:89076507; PMID:3264504
 A:Accession: S02751
 A:Molecule type: DNA
 A:Residues: 82-119 <GH2>
 A:Cross-references: EMBL:M27769
 A:Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown
 R:Grubb, A.; Lofberg, H.
 Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982
 A:Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in the
 A:Reference number: A01270; MUID:82222268; PMID:6283552
 A:Accession: A01270
 A:Molecule type: protein
 A:Residues: 27-131, 'S', 133-146 <GRU>
 R:Ghisso, J.; Jansson, O.; Frangione, B.
 Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986
 A:Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland ty
 A:Reference number: A25434; MUID:86206076; PMID:3517880
 A:Accession: A25434
 A:Molecule type: protein
 A:Residues: 37-93, 'Q', 95-146 <GH1>
 R:Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Bropkin, M.; Borchart, U.; Machleidt, W.
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983
 A:Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystat
 A:Reference number: S01461; MUID:84110053; PMID:6662458
 A:Accession: S12288
 A:Molecule type: protein
 A:Residues: 27-73 <TUR>
 R:Brzin, J.; Popovic, T.; Turk, V.
 Biochem. Biophys. Res. Commun. 118, 103-109, 1984
 A:Title: Human cystatin, a new protein inhibitor of cysteine proteinases.
 A:Accession: A32732
 A:Molecule type: protein
 A:Residues: 27-76 <BR2>
 R:Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jansson, O.; Grubb, A.
 Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
 A:Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst
 A:Reference number: A60552; MUID:90193615; PMID:2315647
 A:Accession: A60552
 A:Molecule type: protein
 A:Residues: 27-49, 'XX', 52-64 <OLA>
 A:Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
 e defective gene is not present in CSF but is found instead in amyloid deposits
 R:Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
 Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
 A:Title: Different forms of human cystatin C.
 A:Reference number: S10607; MUID:91025625; PMID:2222856
 A:Accession: S10607
 A:Molecule type: protein
 A:Residues: 27-53 <POP>
 A:Experimental source: urine, kidney disease
 A:Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also
 R:Grubb, A.; Lofberg, H.; Barrett, A.J.
 FEBS Lett. 170, 370-374, 1984
 A:Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.
 A:Reference number: S01462
 A:Contents: annotation; disulfide bonds
 R:Bertl, P.J.; Storer, A.C.

A:Molecule type: protein
 A:Residues: 28-141 <ISE>
 R:Ramaiahbhu, N.; Reddy, M.S.; Bergey, E.J.; Haraszthy, G.G.; Soni, S.D.; Levine, M.J.
 Biochem. J. 280, 341-352, 1991
 A:Title: Large-scale purification and characterization of the major phosphoproteins and
 A:Reference number: S19279; MUID:92082469; PMID:1747107
 A:Accession: S19279
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 21-55 <RAM>
 C:Comment: Human saliva appears to contain several cysteine proteinase inhibitors that a
 cence. Cystatin SN, with a pI of 7.5, is a much better inhibitor of papain and dipeptidyl
 C:Genetics:
 A:Gene: GDB:CST1
 A:Cross-references: GDB:119815; OMIM:123855
 A:Map position: 20p11.2-20p11.2
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: cysteine proteinase inhibitor; extracellular protein; saliva
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-141/Product: cystatin SA-I #status experimental <MAT>
 F:28-141/Product: cystatin SN #status experimental <MAT>
 F:30-141/Domain: cystatin homology <CYS>
 F:76-80/Region: inhibitory #status predicted
 F:94-104,118-138/Diulfide bonds: #status predicted

Query Match 30.1%; Score 82; DB 1; Length 141;
 Best Local Similarity 35.6%; Pred. No. 0.0028;
 Matches 16; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

OY 2 QYNKESDQKXFRIFRYLVKQVQVTDHLEHYLVNEMQMTTCQKPE 46
 DB 54 EYKATQDYRRPRLRVLRARQQTIVGVNFFPDEVGRTICTKSQ 98

RESULT 11

onchocystatin - nematode (Onchocerca volvulus)
 A:Accession: A43428
 N/Alternate names: cysteine proteinase inhibitor; onchocerciasis antigen
 C/Species: Onchocerca volvulus
 C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
 C/Accession: A43428; B43927
 R:Lustigman, S.; Brozman, B.; Huima, T.; Prince, A.M.; McKerrow, J.H.
 J. Biol. Chem. 267, 17339-17346, 1992
 A:Title: Molecular cloning and characterization of onchocystatin, a cysteine proteinase
 A:Reference number: A43428; MUID:92381053; PMID:1512265
 A:Accession: A43428
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-162 <LUS>
 A:Cross-references: GB:M37105; NID:g159905; PID:g159906
 A:Note: sequence extracted from NCBI backbone (NCBIN:111962, NCBI:P.111963)
 R:Chandraashekar, R.; Maqsood, K.; Alvarez, R.M.; Ogunrinade, A.F.; Lujan, R.; Richards Jr
 J. Clin. Invest. 88, 1460-1466, 1991
 A:Title: Molecular cloning and characterization of recombinant parasite antigens for imm
 A:Reference number: A43927; MUID:92042723; PMID:1840605
 A:Accession: B43927
 A:Molecule type: mRNA
 A:Residues: 'P', 37-57, 'A', 59-71, 'R', 73-83, 'N', 85-126, 'W', 128-162 <CHA>
 A:Cross-references: GB:M60279; NID:g159888
 A:Experimental source: clone OC 9.3
 A:Note: sequence extracted from NCBI backbone (NCBIN:65111, NCBI:P.65113)
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: cysteine proteinase inhibitor

Query Match 29.8%; Score 81; DB 2; Length 162;
 Best Local Similarity 34.1%; Pred. No. 0.0043;
 Matches 14; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

OY 4 NKESDQKXFRIFRYLVKQVQVTDHLEHYLVNEMQMTTCQK 44
 DB 77 NQSDNDEYHMPKILTKVSSQVAGVYKXMDVQVARSQCKK 117

RESULT 12

UDHUP1
 Cystatin S precursor - human
 N/Alternate names: cystatin SA-III; salivary acidic protein-1
 C/Species: Homo sapiens (man)
 C/Date: 25-Feb-1985 #sequence_revision 08-Feb-1996 #text_change 16-Jul-1999
 C/Accession: S17667; S16500; R01272; A29603; S19280; A56608
 R:Bobek, L.A.; Aguirre, A.; Levine, M.J.
 Biochem. J. 278, 627-635, 1991
 A:Title: Human salivary cystatin S. Cloning, sequence analysis, hybridization in situ and
 A:Reference number: S17667; MUID:91378918; PMID:1898352
 A:Accession: S17667
 A:Molecule type: mRNA
 A:Residues: 1-141 <BOB>
 A:Cross-references: EMBL:X54667; NID:g30365; PIDN:CA48478.1; PID:g30366
 R:Laakkin, M.S.; Jensen, U.L.; Setayesh, M.R.; Troxler, R.F.; Oppenheim, F.G.
 Arch. Biochem. Biophys. 288, 664-670, 1991
 A:Title: Salivary cystatin SA-III, a potential precursor of the acquired enamel pellicle
 A:Reference number: S16500; MUID:91378515; PMID:1898055
 A:Accession: S16500
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 21-134, 'D', 136-141 <IHU>
 R:Isemura, S.; Saitoh, E.; Ito, S.; Isemura, M.; Sanada, K.
 J. Biochem. 96, 489-498, 1984
 A:Title: Isolation and amino acid sequence of SP-1, an acidic protein of human whole sal.
 A:Reference number: A91985; MUID:85054716; PMID:6501254
 A:Accession: A01272
 A:Molecule type: protein
 A:Residues: 29-134, 'D', 136-141 <ISE>
 R:Isemura, S.; Saitoh, E.; Ito, S.; Isemura, M.; Sanada, K.
 J. Biochem. 96, 1311-1314, 1984
 A:Title: Cystatin S: a cysteine proteinase inhibitor of human saliva.
 A:Reference number: A91981; MUID:85104877; PMID:6394600
 A:Contents: annotation; inhibitor specificity
 R:Hawke, D.H.; Yuan, P.M.; Wilson, K.J.; Hunkapiller, M.W.
 Biochem. Biophys. Res. Commun. 145, 1248-1253, 1987
 A:Title: Identification of a long form of cystatin from human saliva by rapid microbre
 A:Reference number: A29603; MUID:87270697; PMID:3496880
 A:Accession: A29603
 A:Molecule type: protein
 A:Residues: 21-51 <HAM>
 R:Ramaiahbhu, N.; Reddy, M.S.; Bergey, E.J.; Haraszthy, G.G.; Soni, S.D.; Levine, M.J.
 Biochem. J. 280, 341-352, 1991
 A:Title: Large-scale purification and characterization of the major phosphoproteins and n
 A:Reference number: S19279; MUID:92082469; PMID:1747107
 A:Accession: S19280
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 21-55 <RAM>
 R:Johnson, M.; Richardson, C.F.; Bergey, E.J.; Levine, M.J.; Nancollas, G.H.
 Arch. Oral Biol. 36, 631-636, 1991
 A:Title: The effects of human salivary cystatins and statherin on hydroxyapatite crystal
 A:Reference number: A56608; MUID:92074898; PMID:1741693
 A:Accession: A56608
 A:Molecule type: protein
 A:Residues: 21-36 <UDH>
 A:Note: sequence extracted from NCBI backbone (NCBI:P.67866)
 A:Note: authors designate form without phosphate as cystatin S and form containing one ph
 C:Comment: This protein strongly inhibits papain and ficin, partially inhibits stem brom
 competitively.
 C:Genetics:
 A:Gene: GDB:CST4
 A:Cross-references: GDB:136381
 A:Map position: 20p11.2-20p11.2
 C:Superfamily: cystatin; cystatin homology
 C:Keywords: cysteine proteinase inhibitor; phosphoprotein; saliva
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-141/Product: cystatin S #status predicted <MAT>
 F:30-141/Domain: cystatin homology <CYS>
 F:76-80/Region: inhibitory #status predicted
 F:94-104,118-138/Diulfide bonds: #status predicted

Query Match 29.4%; Score 80; DB 1; Length 141;
 Best Local Similarity 33.3%; Pred. No. 0.005;
 Matches 15; Conservative 12; Mismatches 18; Indels 0; Gaps 0;
 QY 2 QYNKESDDKXHFRIFRYLKVQROVTDHLEVHLNEMQWTTGCKPE 46
 DB 54 EYNKATEDEYRRLPLQVRRARQTFGGVNFDFVEVGRTITKSGQ 98

RESULT 13

T31871
 hypothetical protein F41E6.6 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T31871
 R/Sammons, L.; Murray, J.
 submitted to the EMBL Data Library, July 1997
 A/Description: The sequence of C. elegans cosmid F41E6.
 A/Reference number: Z21095
 A/Accession: T31871
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-498 <SAM>
 A/Cross-references: EMBL:AF016448; PDB:1AAB5956.1; GSPDB:GN00023; CESP:F41E6.6
 A/Experimental source: Strain Bristol N2; clone F41E6
 C/Genetics:
 A/Map position: 5
 A/Intons: 34/3; 102/3; 140/2; 180/2; 336/2; 357/2; 417/1

Query Match 28.3%; Score 77; DB 2; Length 498;
 Best Local Similarity 27.7%; Pred. No. 0.05;
 Matches 13; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 DOYNKESDDKXHFRIFRYLKVQROVTDHLEVHLNEMQWTTGCKPET 47
 DB 42 EKNQSQNDAYKWEIDRTWEYERKLSGINSIFVTLYKIDCKKQGT 88

RESULT 14

JC2040
 cystatin - chum salmon
 N/Alternate names: cysteine proteinase inhibitor
 C/Species: Oncorhynchus keta (chum salmon)
 C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 16-Jul-1999
 C/Accession: JC2040
 R/Koide, Y.; Noso, T.
 Biosci. Biotechnol. Biochem. 58, 164-169, 1994
 A/Title: The complete amino acid sequence of pituitary cystatin from chum salmon.
 A/Reference number: JC2040; MUID:94162738; PMID:7764512
 A/Accession: JC2040
 A/Molecule type: protein
 A/Residues: 1-111 <KOI>
 C/Comment: The intracellular role of this protein is the inhibition of intralysosomal pr
 C/Superfamily: cystatin; cystatin homology
 C/Keywords: cysteine proteinase inhibitor
 F/2-111/Domain: cystatin homology <CYS>
 F/48-52/Region: inhibitory
 F/89-109/Disulfide bonds: #status experimental

Query Match 27.6%; Score 75; DB 1; Length 111;
 Best Local Similarity 30.2%; Pred. No. 0.017;
 Matches 13; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRYLKVQROVTDHLEVHLNEMQWTTGCK 44
 DB 26 EHNKKTNDMFVRQYAKVNAQKQVSGMKYIFTVQMGRTPCRK 68

RESULT 15

JC4918
 cystatin precursor - chum salmon
 C/Species: Oncorhynchus keta (chum salmon)

C/Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jan-2000
 C/Accession: JC4918
 R/Yamashita, M.; Konagaya, S.
 J. Biochem. 120, 483-487, 1996
 A/Title: Molecular cloning and gene expression of chum salmon cystatin.
 A/Reference number: JC4918; MUID:97058285; PMID:8902609
 A/Accession: JC4918
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-132 <YAM>
 A/Cross-references: DDBJ:D86628
 A/Experimental source: liver
 C/Comment: This protein is a homolog of mammalian cystatin C, and acts as an extracellular
 C/Superfamily: cystatin; cystatin homology
 C/Keywords: liver
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/23-132/Domain: cystatin homology <CYS>

Query Match 27.6%; Score 75; DB 2; Length 132;
 Best Local Similarity 30.2%; Pred. No. 0.02;
 Matches 13; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRYLKVQROVTDHLEVHLNEMQWTTGCK 44
 DB 47 EHNKKTNDMFVRQYAKVNAQKQVSGMKYIFTVQMGRTPCRK 89

Search completed: March 18, 2004, 14:23:17
 Job time : 9.43056 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:02:50 ; Search time 5.8333 Seconds
(without alignments)
437.389 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272
Sequence: 1 DQYNKSDDKCHRRIRRVLK.....EYHLNVMQMTCKQKPTTN 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	272	100.0	CS11_HUMAN	Q9H112 homo sapien
2	177	65.1	CS11_MOUSE	O94269 mus musculu
3	127	46.7	CS18_HUMAN	O60676 homo sapien
4	120	44.1	CS18_MOUSE	P32766 mus musculu
5	115	42.3	CS18_RAT	O89969 rattus norv
6	112	41.2	CS18_RAT	P14841 rattus norv
7	110	40.4	CYT2C_MOUSE	O97862 cyrtocolagus
8	108	39.7	CYT2C_MOUSE	P21460 mus musculu
9	105	38.6	CYT2C_MOUSE	O19093 saimiri sci
10	103	37.9	CYT2C_MOUSE	P01038 gallus galli
11	103	37.9	CYT2C_MOUSE	P01034 homo sapien
12	103	37.9	CYT2C_MOUSE	O19092 macaca mula
13	103	37.9	CYT2C_MOUSE	P01035 bos taurus
14	99	36.4	CYT2C_MOUSE	P81061 coturnix co
15	94	34.6	CYT2C_MOUSE	O9H114 homo sapien
16	93	34.2	CYT2C_MOUSE	P09228 homo sapien
17	90	33.1	CYT2C_MOUSE	P08935 bittis ariet
18	89	32.7	CYT2C_MOUSE	P35481 cypripus ca
19	85	31.2	CYT2C_MOUSE	O15828 homo sapien
20	82	30.1	CYT2C_MOUSE	P01037 homo sapien
21	81	29.8	CYT2C_MOUSE	P22085 onchocerca
22	80	29.4	CYT2C_MOUSE	P01036 homo sapien
23	75	27.6	CYT2C_MOUSE	O98967 oncorhynch
24	75	27.6	CYT2C_MOUSE	P81195 oncorhynch
25	74.5	27.4	CYT2C_MOUSE	P28325 homo sapien
26	71	26.1	CYT2C_MOUSE	O76096 homo sapien
27	67	24.6	CYT2C_MOUSE	P19313 rattus norv
28	64	23.5	CYT2C_MOUSE	O89098 mus musculu
29	64	23.5	CYT2C_MOUSE	P01046 bos taurus
30	64	23.5	CYT2C_MOUSE	P01044 bos taurus
31	63	23.0	CYT2C_MOUSE	P29717 candida alb
32	62.5	23.0	CYT2C_MOUSE	P23776 saccharomyc
33	62	22.8	CYT2C_MOUSE	P01047 bos taurus

34	62	22.8	619	1	KNH2_BOVIN	P01045 bos taurus
35	61.5	22.6	137	1	CST9_MOUSE	O92016 mus musculu
36	60	22.1	99	1	CYT_NAJAT	P81714 najia atra (
37	60	22.1	257	1	CG1C_ORYSA	P93411 oryza sativ
38	60	22.1	661	1	KNH_MOUSE	O08677 mus musculu
39	59	21.7	644	1	KNH_HUMAN	P01042 homo sapien
40	57.5	21.1	2200	1	LAR_CAEEL	O9bms caenorhabdi
41	56.5	20.8	483	1	LSH_YEAST	P41921 saccharomyc
42	56	20.6	200	1	SP24_BOVIN	O27967 bos taurus
43	56	20.6	621	1	NIAT_MAIZE	P17571 zea mays (m
44	55.5	20.4	527	1	YBIP_ECOLI	P17585 escherichia
45	55	20.2	525	1	HRG_HUMAN	P04196 homo sapien

ALIGNMENTS

RESULT 1
ID CS11_HUMAN STANDARD; PRT: 137 AA.
AC O9H112; O9H113;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin II precursor.
GN CS11 OR CSTBL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
[1]
SEQUENCE FROM N. A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.B., Connor R.E., Cobby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.B., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leibaeslatho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McKussey A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.B.,
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
"The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:665-871(2001).
-1- SUBCELLULAR LOCATION: Secreted (Potential).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
Isoid=O9H112-1; Sequence=Displayed;
Name=2;
Isoid=O9H112-2; Sequence=VSP_001260;
Note=No experimental confirmation available;
-1- SIMILARITY: Belongs to the cystatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AL096677; CAC13170.1; -
DR EMBL; AL096677; CAC17423.1; -
DR HSSP; P01038; 1490.
DR Gene; HGNC:15959; CST11.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY, 1.
DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
DR Thiol protease inhibitor; Signal; Alternative splicing.
KW SIGNAL 1 25
FT CHAIN 26 137
FT SITE 75 79
FT DISULFID 93 101
FT CARBOHYD 131 134
FT VARSPIC 76 110
SQ SEQUENCE 137 AA; 16375 MW; C5856C39A585C3B CRC64;
Query Match 100.0%; Score 272; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 6,3e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DQYKESDDKXHFRIPLKQVQVTDHLEHNLVEMQWTCQKPEETN 49
Db 52 DQYKESDDKXHFRIPLKQVQVTDHLEHNLVEMQWTCQKPEETN 100

RESULT 2
CS11_MOUSE
ID CS11_MOUSE STANDARD; PRT; 139 AA.

AC Q9D269;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin.11 precursor.
GN CST11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spididymis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,
RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayaishiaki Y.;
RT "Functional annotation of a full-length mouse CDNA collection."
RL Nature 409:685-690(2001).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AK020300; BAB32061.1; -
DR HSSP; P01034; 1696.
DR MED; MGI:1925490; Cst11.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY, 1.
DR PROSITE; PS00287; CYSTATIN; FALSE_NEG.
DR Thiol protease inhibitor; Signal.
KW SIGNAL 1 28
FT CHAIN 29 139
FT SITE 76 80
FT DISULFID 94 102
FT CARBOHYD 134 135
SQ SEQUENCE 139 AA; 16217 MW; F228D9815FA32640 CRC64;
Query Match 65.1%; Score 177; DB 1; Length 139;
Best Local Similarity 64.6%; Pred. No. 9,4e-16;
Matches 31; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DQYKESDDKXHFRIPLKQVQVTDHLEHNLVEMQWTCQKPEET 48
Db 53 DQYKESDDKXHFRIPLKQVQVTDHLEHNLVEMQWTCQKPEET 100

RESULT 3
CST8_HUMAN
ID CST8_HUMAN STANDARD; PRT; 142 AA.

AC O60676;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
DE 8).
GN CST8 OR CRBS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=95344753; PubMed=7619504;
RA Cornwall G.A., Hann S.R.,
RT "Transient appearance of CRBS protein during spermatogenesis and
RT caput epididymal sperm maturation."
RL Mol. Reprod. Dev. 41:37-46(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dham P.D., Dunn M.,
RA Ellington D.V., Frankland J.A., Fraser A., French L., Garner P.,
RA Graiham G., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehesvahti M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McCormachie L.J., McIlroy K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramay H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showgreen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 CC EMBL: AF059244; AAC14707.1; -;
 CC EMBL: AL109954; CAB64234.1; -;
 CC HSSP: P01034; 1G96.
 DR GeneW: HNCN:2480; CST8.
 DR GO: GO:0004869; F:cysteine protease inhibitor activity; TAS.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR Thiol protease inhibitor; Signal; Polymorphism.
 KW Thiol protease inhibitor; Signal; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 142
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 95 105 BY SIMILARITY.
 FT DISULFID 119 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 142 142 A -> P (in dbSNP:1054633).
 FT SIGNAL 142 142 /FTID=VAR 014527.
 SQ SEQUENCE 142 AA; 16275 MW; 9A3512757E0F4ECD CRC64;
 Query Match 46.7%; Score 127; DB 1; Length 142;
 Best Local Similarity 50.0%; Pred. No. 2.5e-09;
 Matches 24; Conservative 12; Mismatches 12; Indels 0; Gaps 0;
 QY 2 QYNKESDQKTHFRFLVAKVQROVTDHLEHLNEMQWTTQCKPRTN 49
 Db 55 EYKNESEDQKVFVAVKTIQAOQVNTLELYLIDVEIARSDCKRLSTN 102
 RESULT 4
 CST8_MOUSE STANDARD; PRT; 142 AA.
 AC P32766; O89102; (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-related epididymal spermato-genic protein precursor (Cystatin-
 DE related epididymal specific protein) (Cystatin 8).
 GN CST8 OR CRES.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;

RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CH, and CD-1;
 RX MEDLINE=9247899; PubMed=10229662;
 RA Cornwall G.A., Hsia N., Sutton H.G.;
 RT "Structure, alternative splicing and chromosomal localization of the
 RT cystatin-related epididymal spermato-genic gene";
 RL Biochem. J. 340:85-93(1999).
 RN [2]
 RN SEQUENCE OF 4-142 FROM N.A.
 RP TISSUE=Epididymis;
 RX MEDLINE=93078799; PubMed=1280328;
 RA Cornwall G.A., Orgebin-Crist M.-C., Hann S.R.;
 RT "The CRES gene: a unique testis-regulated gene related to the cystatin
 RT family is highly restricted in its expression to the proximal region
 RT of the mouse epididymus";
 RL Mol. Endocrinol. 6:1653-1664(1992).
 CC -1- FUNCTION: Performs a specialized role during sperm development and
 CC maturation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Proximal caput region of the epididymis. Lower
 CC expression in the testis. Within the testis it is localized to the
 CC elongating spermatids, whereas within the epididymis it is
 CC exclusively synthesized by the proximal caput epithelium.
 CC -1- INDUCTION: Testicular factors or hormones other than androgens
 CC present in the testicular fluid may be involved in the regulation
 CC of CRES gene expression.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 CC EMBL: AF091503; AAC61754.1; -;
 CC EMBL: AF090691; AAC36316.1; -;
 CC EMBL: S49261; AAC35390.1; -;
 CC PIR: A45361; A45361.
 DR HSSP: P01034; 1G96.
 DR MGD: MGI:107161; Cst8.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; cystatin; 1.
 DR SMART: SM00043; CY; 1.
 KW Thiol protease inhibitor; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 142
 FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 95 105 BY SIMILARITY.
 FT DISULFID 119 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 4 15 PMLSILFPIIP -> GTREQVGSBSK (IN REF. 2).
 SQ SEQUENCE 142 AA; 16288 MW; 50B446B9F66673E CRC64;
 Query Match 44.1%; Score 120; DB 1; Length 142;
 Best Local Similarity 43.2%; Pred. No. 1.9e-08;
 Matches 19; Conservative 14; Mismatches 11; Indels 0; Gaps 0;
 QY 2 QYNKESDQKTHFRFLVAKVQROVTDHLEHLNEMQWTTQCKP 45
 Db 55 EYKNESEDQKVFVAVKTIHAKLQITDMEYQIDVOISRSNCKP 98
 RESULT 5
 CST8_RAT STANDARD; PRT; 142 AA.
 AC O88969;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cystatin-related epididymal spermatogenic protein precursor (Cystatin
DE 8).
GN CSTR OR CRES.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Epididymis;
RX MEDLINE=99247899; PubMed=10229662;
RA Cornwall G.A., Hsia N., Sutton H.G.;
RT "Structure, alternative splicing and chromosomal localization of the
RT cystatin-related epididymal spermatogenic gene.";
RL Biochem. J. 340:85-93(1999).
CC -1- FUNCTION: Performs a specialized role during sperm development and
CC maturation.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF090692; AAC36317.1; -.
DR HSBP; P01034; I696.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 142 CYSTATIN-RELATED EPIDIDYMAL SPERMATOGENIC
FT PROTEIN.
FT SITE 77 81 SECONDARY AREA OF CONTACT (POTENTIAL).
FT DISULFID 95 105 BY SIMILARITY.
FT DISULFID 119 139 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ SEQUENCE 142 AA; 16246 MW; F8B73F5A6B6CB34 CRC64;

Query March 42.3%; Score 115; DB 1; Length 142;
Best Local Similarity 43.2%; Pred. No. 8.5e-08;
Matches 19; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

OY 2 QYNKESDDKXHFRIFRVLKVRQVTDHLEHLNEMQMTTCQKP 45
Db 55 EYNGSEDKYFLDLKTLHATLQITDREWHYIDVQISRSNCRKP 98

RESULT 6
ID CYTC_RAT STANDARD; PRT; 127 AA.
CYTC_RAT
AC P14841;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Fragment).
GN CSTR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Buffalo;
RX MEDLINE=90092122; PubMed=2689174;
RA Cole T., Dickson P.W., Esnard F., Averill F., Risbridger G.,
RA Gauthier F., Schneider G.;
RT "The CDNA structure and expression analysis of the genes for the

```

```

RT cysteine proteinase inhibitor cystatin C and for beta 2-microglobulin
RT in rat brain.";
RL Eur. J. Biochem. 186:35-42(1989).
RN [2]
RP SEQUENCE OF 8-127.
RX MEDLINE=90380276; PubMed=2400577;
RA Esnard F., Esnard A., Faucher D., Capony J.-P., Derancourt J.,
RA Brillard M., Gauthier F.;
RT "Rat cystatin C: the complete amino acid sequence reveals a site for
RT N-glycosylation.";
RL Biol. Chem. Hoppe-Seyler 371:161-166(1990).
RN [3]
RP SEQUENCE OF 8-49.
RX MEDLINE=88313020; PubMed=3044831;
RA Esnard A., Esnard F., Faucher D., Gauthier F.;
RT "Two rat homologues of human cystatin C.";
RL FEBS Lett. 236:475-478(1988).
RN [4]
RP SEQUENCE OF 8-20.
RC TISSUE=Sertoli cells;
RX MEDLINE=92225121; PubMed=1563513;
RA Esnard A., Esnard F., Guillon F., Gauthier F.;
RT "Production of the cysteine proteinase inhibitor cystatin C by rat
RT Sertoli cells.";
RL FEBS Lett. 300:131-135(1992).
CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity. Known to inhibit cathepsin B,
CC H. and L.
CC -----
CC -1- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16957; CAA34831.1; -.
DR PIR; S07085; S07085.
DR PIR; S10587; S10587.
DR HSBP; P01034; I696.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Signal.
FT NON_TER 1 1
FT SIGNAL 8 127 CYSTATIN C.
FT CHAIN 18 18 REACTIVE SITE.
FT ACT SITE 62 66 SECONDARY AREA OF CONTACT.
FT SITE 80 90 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 125 25 A -> E (IN REF. 2).
FT CONFLICT 25 25
SQ SEQUENCE 127 AA; 14039 MW; 78F70158B7325853 CRC64;

Query March 41.2%; Score 112; DB 1; Length 127;
Best Local Similarity 43.5%; Pred. No. 1.8e-07;
Matches 20; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

OY 2 QYNKESDDKXHFRIFRVLKVRQVTDHLEHLNEMQMTTCQKPE 47
Db 40 EYNGSNDAYHSRAIQVVRARQVLVAGINYYLDVEMRTTCTISQT 85

RESULT 7
ID CYTC_RABIT STANDARD; PRT; 148 AA.
CYTC_RABIT
AC O97862;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=129/SV; TISSUE=Liver;
RX  MEDLINE=95137392; PubMed=7835704;
RT  Hun C., Nagle J.W., Kozak C.A., Abrahamson M., Karlsson S.;
RA  "Structural organization, expression and chromosomal mapping of the
RL  mouse cystatin-C-encoding gene (Cst3).";
RN  Gene 152:221-226(1995).
RP  [3]
RC  SEQUENCE FROM N.A.
RX  STRAIN=ILS, and ISS;
RT  MEDLINE=21363810; PubMed=11471062;
RA  Ehlinger M.A., Thompson J., Conroy O., Xu Y., Yang F., Caniff J.,
RT  Beeson M., Gordon L., Bennett B., Johnson T.E., Sikele J.M.;
RT  "High-throughput sequence identification of gene coding variants
RL  within alcohol-related QTLs.";
RN  Mamm. Genome 12:657-663(2001).
RP  [4]
RC  SEQUENCE FROM N.A.
RX  MEDLINE=22388257; PubMed=12477932;
RT  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  DiChencko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA  Bonak S.A., McGowan P.J., McKern K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Weale A.C., Haines S., Garcia A.M., Gay L.J., Hulik S.W.,
RA  Vallano D.K., Muzny D.M., Sodergren R.J., Lu X., Gibbs S.A., Sanchez
RA  Fahy J., Helton E., Kettelman M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA  Butcherfield J.S.N., Krzyzanski M.I., Skaleka U., Smallus D.E.,
RA  Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length
RL  human and mouse cDNA sequences.";
RT  Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
CC  -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC  thought to serve an important physiological role as a local
CC  regulator of this enzyme activity.
CC  -1- SIMILARITY: Belongs to the cystatin family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M59470; AAA63298.1; -
DR  EMBL; U10098; AAB41056.1; -
DR  EMBL; AF483486; AAL50760.1; -
DR  EMBL; AF483487; AAL50761.1; -
DR  EMBL; BC002072; AAL02072.1; -
DR  PIR; A36163; A36163.
DR  HSSP; P01034; 1G96.
DR  MGD; MG1102519; Cst3.
DR  InterPro; IPR000010; Cystatin.
DR  Pfam; PF00031; Cystatin; 1.
DR  SMART; SMO0043; CY; 1.
DR  PROSITE; PS00287; CYSTATIN; 1.
KW  Thiol protease inhibitor; Signal.
FT  SIGNAL 1 20
FT  CHAIN 1 140 CYSTATIN C.
FT  ACT_SITE 31 31 REACTIVE SITE.
FT  SITE 75 79 SECONDARY AREA OF CONTACT.
FT  DISULFID 93 103 BY SIMILARITY.
FT  DISULFID 117 137 BY SIMILARITY.
FT  CONFLICT 16 16 A -> G (IN REF. 1).
FT  CONFLICT 84 84 L -> F (IN REF. 1).

```

```

SQ SEQUENCE 140 AA; 15531 MW; 3A563406DD58D0F5 CRC64;

Query Match      39.7%; Score 108; DB 1; Length 140;
Best Local Similarity 43.5%; Pred. No. 6.6e-07;
Matches 20; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHPRIFRVLKVGVVDHLDEHYLVNEMOMTTGCKPPT 47
   :|::|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
Db 53 EYNKASNDWYHSRALQVVARAKQLAVAGVNYFLDVEMGTTCRKQT 98

RESULT 9
CYTC_SAISC STANDARD; PRT; 146 AA.
AC O19053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor...
GN CST3.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euarchonta; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_Taxid=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker J.C.; Levy E.;
RT "Cystatin C. Icelandic-like mutation in an animal model of
RL cerebrovascular beta-amyloidosis.";
RT Stroke 27:2080-2085(1996).
CC -! FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC thought to serve an important physiological role as a local
CC regulator of this enzyme activity.
CC -! SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52028; AAB64051.1; -.
DR HSP; P01034; IGS6.
DR InterPro; IPRO00010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 146 CYSTATIN C.
FT ACT SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFD 99 109 BY SIMILARITY.
FT DISULFD 123 143 BY SIMILARITY.
SQ SEQUENCE 146 AA; 15946 MW; 08196353C030EA3 CRC64;

Query Match      38.6%; Score 105; DB 1; Length 146;
Best Local Similarity 44.2%; Pred. No. 1.7e-06;
Matches 19; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHPRIFRVLKVGVVDHLDEHYLVNEMOMTTGCKPPT 44
   :|::|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
Db 59 EYNKASNDWYHSRALQVVARAKQLAVAGVNYFLDVEMGTTCRKQT 101

RESULT 10
CYT_CHICK STANDARD; PRT; 139 AA.
ID_CYT_CHICK ID_P01038;
DT 21-JUL-1986 (Rel. 01, Created)
```

DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin precursor (Egg-white cystatin).
OS Gallus gallus (Chicken).
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CX NCBI_TaxID=9031.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90008873; PubMed=2793849;
RA Coletta R., Sakaguchi Y., Nagase H., Bird J.W.C.;
RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,
RT and tissue distribution."
J Biol. Chem. 264:17164-17169(1989).
RN [2]
RP SEQUENCE OF 24-139.
RX MEDLINE=94178305; PubMed=6712597;
RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;
RT "Cystatin. Amino acid sequence and possible secondary structure."
J Biochem. J. 217:813-817(1984).
RN [3]
RP SEQUENCE OF 24-139.
RX MEDLINE=94110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
RA Machleidt W.;
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
RT of cystatin from chicken egg white."
Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [4]
RP CHARACTERIZATION OF PROTEIN.
RX MEDLINE=83256421; PubMed=6409085;
RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,
RA Sunter D.C., Barrett A.J.;
RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved
RT purification from egg white, characterization, and detection in
RT chicken serum."
Biochem. J. 211:129-138(1983).
RN [5]
RP DISULFIDE BONDS.
RX Grubb A., Loeffberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
RT cystatin."
FEBS Lett. 170:370-374(1984).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=89252033; PubMed=2721673;
RA Laber B., Krieglstein K., Henschel A., Kos J., Turk V., Huber R.,
RA Bode W.;
RT "The cysteine proteinase inhibitor chicken cystatin is a
RT phosphoprotein."
FEBS Lett. 248:162-168(1989).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=89052676; PubMed=3191914;
RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshkov A.,
RA Brzin J., Kos J., Turk V.;
RT "The 2.0 A X-ray crystal structure of chicken egg white cystatin and
RT its possible mode of interaction with cysteine proteinases."
EMBO J. 7:2593-2599(1988).
RN [8]
RP STRUCTURE BY NMR.
RX MEDLINE=94087719; PubMed=8263912;
RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,
RA Auerwald E.A., Jeanicke R., Oeschkinat H.;
RT "The structures of native phosphorylated chicken cystatin and of a
RT recombinant unphosphorylated variant in solution."
J. Mol. Biol. 234:1048-1059(1993).
CC -I- FUNCTION: This protein binds tightly to and inhibits a variety of
CC thiol proteases including ficin, papain, and cathepsins B, C, H,
CC and L. Although isolated from egg white, it is also present in
CC serum.
CC -I- SIMILARITY: Belongs to the cystatin family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-ebi.ch/announce/>
 CC or send an email to license@isb-ebi.ch).
 CC -----
 DR EMBL: J05077; AAA48744.1; -
 DR PIR: A34456; UDCH.
 DR PDB: 1AE7; 31-JAN-94.
 DR PDB: 1A67; 27-MAY-98.
 DR PDB: 1A90; 17-JUN-98.
 DR InterPro: IPR000010; Cystecatin.
 DR Pfam: PF00031; cystecatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 KM Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.
 FT SIGNAL 1..23
 FT CHAIN 24..139
 FT ACT SITE 32..32
 FT SITE 76..80
 FT DISULFID 94..104
 FT DISULFID 118..138
 FT MOD_RES 103..103
 FT STRAND 35..36
 FT TURN 39..40
 FT HELIX 42..51
 FT TURN 52..52
 FT HELIX 53..56
 FT TURN 57..58
 FT STRAND 63..77
 FT STRAND 81..95
 FT TURN 96..97
 FT TURN 99..100
 FT HELIX 101..108
 FT STRAND 115..125
 FT TURN 126..129
 FT STRAND 130..139
 SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;
 Query Match 37.94; Score 103; DB 1; Length 139;
 Best Local Similarity 44.24; Pred. No. 2.9e-06;
 Matches 19; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 QY 2 QYNKESDDKHYFRIYFLVAKVQGVTDHLEVHLNVMQWTCOK 44
 :||:|||||:|||||:|||||:|||||:|||||
 54 EYNRASNDKYSRVRVYISAKRLQVSGIKTILVEIGRTTCPK 96
 Db
 RESULT 11
 CYTC_HUMAN STANDARD; PRT; 146 AA.
 AC P01034;
 DT 21-UTR-1986 (Rel. 01, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cystecatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
 DE (Post-gamma-globulin).
 GN CSTR3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=87219149; PubMed=3495457;
 RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;
 RA "Molecular cloning and sequence analysis of cDNA coding for the
 RA precursor of the human cysteine proteinase inhibitor cystecatin C.";
 RT FEBS Lett. 216:229-233(1987).
 RL

 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukocyte;
 RX MEDLINE=90303202; PubMed=2363674;
 RA Abrahamson M., Olafsson I., Palsdottir A., Ulfvack M., Lundwall A.,
 RA Jansson O., Grubb A.;
 RT "Structure and expression of the human cystecatin C gene.";
 RL Biochem. J. 268:287-294(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (HCHWA VARIANT).
 RC TISSUE=Brain;
 RX MEDLINE=89235594; PubMed=2541223;
 RA Levy E., Lopez-Otin C., Ghiso J., Geltner D., Frangione B.;
 RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is
 RT related to a mutation in the cystecatin C gene, an inhibitor of
 RT cysteine proteases.";
 RL J. Exp. Med. 169:1771-1778(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89350949; PubMed=2764935;
 RA Saitoh E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,
 RA Iseura S., Sanada K.;
 RT "The human cystecatin C gene (CSTR3) is a member of the cystecatin gene
 RT family which is localized on chromosome 20.";
 RL Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;
 RT "Acquisition of complex patterns of differential expression in
 RT epithelial cell populations during the evolution of type 2 cystecatin
 RT genes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.B., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehesvaho M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McComachie L.J., McElay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Rose M.T., Scott C.B., Sehra H.K., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wali M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whiteaker P., Willey D.L., Williams L., Williams S.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S.A., McGee K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting J., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallue D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 27-146.
 RX MEDLINE=8222268; PubMed=6283552;
 RA Grubb A., Loeffberg H.,
 RT "Human gamma-trace, a basic microprotein: amino acid sequence and
 RT presence in the adenohypophysis.",
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).
 RN [9]
 RP SEQUENCE OF 27-73.
 RX MEDLINE=84110059; PubMed=6662498;
 RA Turk V., Brzin J., Longer M., Riconja A., Eropkin M., Borchart U.,
 RA Machleidt W.,
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
 RT of cystatin from chicken egg white.",
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1993).
 RN [10]
 RP SEQUENCE OF 27-76.
 RX MEDLINE=84128015; PubMed=6365094;
 RA Brzin J., Popovic T., Turk V.,
 RT "Human cystatin, a new protein inhibitor of cysteine proteinases.",
 RL Biochem. Biophys. Res. Commun. 118:103-109(1984).
 RN [11]
 RP DISULFIDE BONDS.
 RA Grubb A., Loeffberg H., Barrett A.J.,
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
 RT cystatin.",
 RL FEBS Lett. 170:370-374(1984).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.
 RX MEDLINE=21173909; PubMed=11276250;
 RA Janowski R., Kozak M., Janowska E., Grzonka Z., Grubb A.,
 RA Abramson M., Jaskolski M.,
 RT "Human cystatin C, an amyloidogenic protein, dimerizes through
 RT three-dimensional domain swapping.",
 RL Nat. Struct. Biol. 8:316-320(2001).
 RN [13]
 RP VARIANT GLN-94.
 RX MEDLINE=92316504; PubMed=1352269;
 RA Abramson M., Jonasdottir S., Olafsson I., Jensen O., Grubb A.,
 RT "Hereditary cystatin C amyloid angiopathy: identification of the
 RT disease-causing mutation and specific diagnosis by polymerase chain
 RT reaction based analysis.",
 RL Hum. Genet. 89:377-380(1992).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SUBUNIT: Homodimer.
 CC -1- TISSUE SPECIFICITY: Expressed in highest levels in the epididymis,
 CC vas deferens, brain, thymus, and ovary and the lowest in the
 CC submandibular gland.
 CC -1- DISEASE: Defects in CSN3 are a cause of hereditary cerebral
 CC hemorrhage with amyloidosis (HCHWA) [MIM:105150] also known as
 CC cerebral amyloid angiopathy (CAA) or cerebroarterial amyloidosis
 CC Icelandic type. HCHWA is characterized by a thickening of the
 CC cerebral arteries walls with deposition of material with the
 CC characteristics of amyloid.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: X05607; CAA29096.1; -;
 DR EMBL: X52255; CAA36497.1; -;
 DR EMBL: M27891; AAA52164.1; -;
 DR EMBL: M27889; AAA52164.1; JOINED.
 DR EMBL: M27890; AAA52164.1; JOINED.
 DR EMBL: X61681; CAA43856.2; -;
 DR EMBL: X61682; CAA43856.2; JOINED.
 DR EMBL: X61683; CAA43856.2; JOINED.
 DR EMBL: AF319564; AAK11570.1; -;
 DR EMBL: AL121894; CAC05424.1; -;
 DR EMBL: BC013083; AAI13083.1; -;
 DR PIR: S10216; UDHU.
 DR PDB: 1G96; 06-APR-01.
 DR Genew; HGNC:2475; CST3.
 DR MIM; 604312; -;
 DR MIM; 105150; -;
 DR InterPro: IPR000010; Cystatin.
 DR Pfam; PF00093; Cyt. 1.
 DR SMART; SM00043; Cyt. 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR TM01; TM01; Cystatin; 1.
 KM Polymorphism; 3D-structure.
 FT SIGNAL 1 26
 FT CHAIN 27 146 CYSTATIN C.
 FT ACT_SITE 37 37 REACTIVE SITE.
 FT SITE 81 85 SECONDARY AREA OF CONTACT.
 FT DISULFID 99 109
 FT DISULFID 123 143
 FT
 Query Match 37.9%; Score 103; DB 1; Length 146;
 Best Local Similarity 40.0%; Pred. No. 3e-06;
 Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
 Oy 2 QYNESDDKHYFRIFRYLKVQVQTHLEHYLVNEMQWTCORPE 46
 Db 59 EYKASNDMYHSRALQVVRARQIVAGVNFVDFVLEGLRTTCRTQ 103
 RESULT 12
 CYTC_MACMU STANDARD; PRT; 146 AA.
 AC 019092;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor.
 GN CSN3.
 OS Macaca mulatta (Rhesus macaque).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecinae; Macaca.
 CC NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97054523; PubMed=8898820;
 RA Wei L.H., Walker L.C., Levy E.,
 RT "Cystatin C. Icelandic-like mutation in an animal model of
 RT cerebrovascular beta-amyloidosis.",
 RL Stroke 27:2080-2085(1996).
 CC -1- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
 CC thought to serve an important physiological role as a local
 CC regulator of this enzyme activity.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U51912; AAB64050.1; -.
 DR HSSP; P01034; 1G96.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR Thiol protease inhibitor; Amyloid; signal.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 146 CYSTATIN C.
 FT ACT SITE 37 37 REACTIVE SITE.
 FT SITE 81 85 SECONDARY AREA OF CONTACT.
 FT DISULFID 99 109 BY SIMILARITY.
 FT DISULFID 123 143 BY SIMILARITY.
 SQ SEQUENCE 146 AA; 15857 MW; F0B3B8774A29DF26 CRC64;

Query Match 37.9%; Score 103; DB 1; Length 146;
 Best Local Similarity 40.0%; Pred. No. 3e-06;
 Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRFVLYKQOVTDHLEHLNEMQWTCQKPE 46
 Db 59 EYNKASNDMTHSRALQVVRARAKQIVAGVNFVLDVLELGRITCTKIQ 103

RESULT 13
 CYTC_BOVIN STANDARD; PRT; 148 AA.
 AC P01035;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-JUL-1998 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin C precursor (Colostrum thiol proteinase inhibitor).
 GN CST3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.
 RC TISSUE=Cerebrospinal fluid, and Choroid plexus;
 RX MEDLINE=98094199; PubMed=9434110;
 RA Olsson S.-L., Ek B., Wilm M., Broberg S., Rask L., Bioeek I.;
 RT "Molecular cloning and N-terminal analysis of bovine cystatin C
 RT identification of a full-length N-terminal region.";
 RL Biochim. Biophys. Acta 1343:203-210(1997).
 RN [2]
 RP SEQUENCE OF 37-148.
 RX MEDLINE=85231205; PubMed=3891407;
 RA Hlarado M., Tsumasawa S., Sakiyama F., Ninobe M., Fujii S.;
 RT "Complete amino acid sequence of bovine colostrum low-Mr cysteine
 RT proteinase inhibitor.";
 RL FBS Lett. 186:41-45(1985).
 CC -1- FUNCTION: This is a thiol proteinase inhibitor.
 CC -1- MASS SPECTROMETRY: MW=13420; METHOD=MALDI.
 CC -1- SIMILARITY: Belongs to the cystatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Y10811; CAAT1771.1; -.
 DR HSSP; P01034; 1G96.
 DR InterPro; IPR000010; Cystatin.

DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR Thiol protease inhibitor; Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 30 PROBABLE.
 FT CHAIN 31 148 CYSTATIN C.
 FT MOD RES 31 31 PYROGLUTAMATE CARBOXYLIC ACID (PROBABLE).
 FT ACT SITE 40 40 REACTIVE SITE.
 FT SITE 84 88 SECONDARY AREA OF CONTACT.
 FT DISULFID 102 112 BY SIMILARITY.
 FT DISULFID 126 146 BY SIMILARITY.
 SQ SEQUENCE 148 AA; 16265 MW; E8740F837CEB9F0E CRC64;

Query Match 37.9%; Score 103; DB 1; Length 148;
 Best Local Similarity 40.0%; Pred. No. 3.1e-06;
 Matches 18; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRFVLYKQOVTDHLEHLNEMQWTCQKPE 46
 Db 62 EFNKRSNDAYQSRVVRVVRARAKQVSGMNFVLDVLELGRITCTKTSQ 106

RESULT 14
 CYT_COTUA STANDARD; PRT; 116 AA.
 AC P81061;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin (Egg-white cystatin).
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Coturnix.
 NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Egg white; PubMed=9276465;
 RX Gerhardt B., Eckerstrom C., Torguato R.,
 RA Wiltman J., Kolb H.J., Machleidt W., Fritz H., Auerwald E.A.;
 RT "Quail cystatin: isolation and characterization of a new member of
 RT the cystatin family and its hypothetical interaction with cathepsin
 RT B.";
 RL FBS Lett. 412:551-558(1997).
 CC -1- FUNCTION: This protein binds tightly to and inhibits papain and
 CC cathepsin B.
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC HSSP; P01036; 1CEW.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR Thiol protease inhibitor; Phosphorylation.
 FT ACT SITE 9 9 REACTIVE SITE.
 FT SITE 53 57 SECONDARY AREA OF CONTACT.
 FT DISULFID 71 81
 FT DISULFID 95 115
 FT MOD RES 80 80
 SQ SEQUENCE 116 AA; 13093 MW; 48248621053A2PF0 CRC64;

Query Match 36.4%; Score 99; DB 1; Length 116;
 Best Local Similarity 39.5%; Pred. No. 7.6e-06;
 Matches 17; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 2 QYNKESDDKYHFRFVLYKQOVTDHLEHLNEMQWTCQK 44
 Db 31 EYNKASNDKTSKRVRIISAKQQLVSGIKIMEVIGRTTCKP 73

RESULT 15
 CSTL_HUMAN STANDARD; PRT; 165 AA.

AC 09H114;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cystatin-like 1 precursor.
 GN CSTL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafton D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leivaeslath M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vautin M., Wall M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams D.R., Beck S.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:665-871(2001).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the cystatin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; AL096677; CAC03440.2; -
 DR HSSP; P01038; I190.
 DR Genew; HGNC:15958; CSTL1.
 DR InterPro; IPR000010; Cystatin.
 DR PROSITE; PS00287; CYSTATIN; FALSE NEG.
 KM Thiol protease inhibitor; Signal_
 FT STGNAL 1 19 POTENTIAL.
 FT CHAIN 20 165 CYSTATIN-LIKE 1.
 FT SITE 93 97 SECONDARY AREA OF CONTACT (POTENTIAL).
 FT DISULFID 111 121 BY SIMILARITY.
 FT DISULFID 134 154 BY SIMILARITY.
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 165 AA; 19312 MW; 9D6D685875DAE8A CRC64;

Query Match 34.6%; Score 94; DB 1; Length 165;
 Best Local Similarity 28.4%; Pred. No. 4.9e-05;

Matches 19; Conservative 15; Mismatches 13; Indels 20; Gaps 1;
 QY 3 YNKSDPKYHFRIFRYLKVQ-----ROYTDHLEHIANVEMQWTTTC 42
 DB 52 YNNASNDTYIYRQRLIRSGMQRVSHMGLGVHTNTTDSRQLTGVETVTKIGWTKC 111
 QY 43 QKPEPTN 49
 DB 112 KRNDTSN 118
 Search completed: March 18, 2004, 14:16:12
 Job time : 5.83333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 14:05:55 ; Search time 29.3611 Seconds
(without alignments) 526.560 Million cell updates/sec

Title: US-09-941-314-13

Perfect score: 272
Sequence: 1 DQYKESDDKXHPRIFRVLK.....EYHLNVMQWTTQCKPPTTN 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaeoplasmid:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	272	100.0	138	4	Q8WU6
2	162	59.6	139	11	Q8K5A3
3	128	47.1	103	4	Q8WU5
4	122	44.9	141	11	Q9DAP1
5	122	44.9	141	11	Q80ZNS
6	108	39.7	140	11	Q9BFX9
7	97	35.7	112	13	Q9BSR4
8	97	35.7	112	13	Q9BSR3
9	84	30.9	140	11	Q80Y72
10	83	30.5	425	3	Q12700
11	81	29.8	109	5	Q9TY65
12	81	29.8	148	5	Q9NH95
13	80	29.4	161	5	Q16159
14	79	29.0	146	11	Q8K397
15	79	29.0	149	11	Q9DIB1
16	78	28.7	149	11	Q8VHC1

17	77	28.3	498	5	Q16454
18	74	27.2	144	13	Q8JFUS
19	74	27.2	148	11	Q9JMR4
20	72	26.5	125	5	Q25620
21	71	26.1	167	4	Q72408
22	68	25.0	157	5	Q17108
23	67	24.6	127	5	P90698
24	67	24.6	284	16	Q88207
25	64.5	23.7	443	5	Q8ICX8
26	64	23.5	81	6	Q29212
27	64	23.5	167	11	Q9QWL5
28	63.5	23.3	787	11	Q8B182
29	63	23.2	438	1	Q9URJ8
30	62.5	23.0	302	4	Q9H740
31	62.5	23.0	421	4	Q9NXS0
32	62.5	23.0	995	4	Q9HC18
33	62	22.8	693	11	Q8K145
34	61.5	22.6	587	9	Q859J1
35	61.5	22.6	587	9	Q859J0
36	61	22.4	159	4	Q8TD53
37	60.5	22.2	133	5	Q8WVB6
38	60	22.1	204	4	Q8TC7
39	60	22.1	275	4	Q8WYG2
40	60	22.1	357	4	Q72465
41	60	22.1	462	13	Q72Y91
42	60	22.1	462	13	Q75YH2
43	60	22.1	465	13	Q801E5
44	59.5	21.9	128	11	Q9DAN8
45	59.5	21.9	586	5	Q94510

ALIGNMENTS

RESULT 1	Q8WU6	PRELIMINARY;	PRT;	138 AA.
AC	Q8WU6	Q8WU6		
DT	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)			
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)			
DE	SC13.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamil K.G., Liu O., Zhang Y.-L., French F.S., Hall S.H.;			
RT	"SC13: A novel epididymal specific member of the cystatin family."			
RL	Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF335480; AAL71991.1; -			
DR	GO; GO:0004869; F:Cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR000010; Cystatin.			
DR	Pfam; PF00003; Cystatin; 1.			
DR	SMART; SM00043; CY; 1.			
DR	SEQUENCE 138 AA; 16506 MW; E49440ACA3585C64 CRC64;			
SO	Q8WU6			
Query Match	100.0%; Score 272; DB 4; Length 138;			
Best Local Similarity	100.0%; Pred. No. 7.2e-28;			
Matches	49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	1 DQYKESDDKXHPRIFRVLKQROVTDHLEHLNVMQWTTQCKPPTTN 49			
DB	53 DQYKESDDKXHPRIFRVLKQROVTDHLEHLNVMQWTTQCKPPTTN 101			
RESULT 2	Q8K5A3	PRELIMINARY;	PRT;	139 AA.
ID	Q8K5A3			
AC	Q8K5A3			
DT	01-OCT-2002 (TREMBlrel. 22, Created)			
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)			

```

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cystatin 11.
GN CS711.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley;
RA Hamil K.G., Hall S.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF501290; AAM21709.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 139 AA; 1686 MW; E1E36DB786B4D08C CRC64;

Query Match 59.6%; Score 162; DB 11; Length 139;
Best Local Similarity 56.5%; Pred. No. 2e-13;
Matches 26; Conservative 14; Mismatches 6; Indels 0; Gaps 0;

QY 1 DQYKESDDKXHFRIFRVLKQVQVTDHLEYNLVEMQWTTCCOK 46
DB 53 EYVKKSEDDLYNFRILIKLEKQWTHMEPHITVEMQRTTCLKTE 98

RESULT 3
Q8MXU5 PRELIMINARY; PRT; 103 AA.
ID Q8MXU5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SC13delta.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamil K.G., Liu Q., Zhang Y.-L., French F.S., Hall S.H.;
RT "SC13: A novel epididymal specific member of the cystatin family.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335481; AAL71992.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 103 AA; 1285 MW; 05DD92C47387B022 CRC64;

Query Match 47.1%; Score 128; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DQYKESDDKXHFRIFRVLKQVQ 24
DB 53 DQYKESDDKXHFRIFRVLKQVQ 76

RESULT 4
Q9DAP1 PRELIMINARY; PRT; 141 AA.
ID Q9DAP1;
AC Q9DAP1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 1700006C19R1k protein.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konno S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staudl F., Suzuki R., Tomita M., Wagner U., Maehiro T.,
RA Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005665; BAB24175.1; -
DR HSSP; P01038; ICEW.
DR MGD; MGI:1916544; 1700006C19R1k.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 141 AA; 1681 MW; C20FA0D8B1AC378C CRC64;

Query Match 44.9%; Score 122; DB 11; Length 141;
Best Local Similarity 51.2%; Pred. No. 3.6e-08;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYKESDDKXHFRIFRVLKQVQVTDHLEYNLVEMQWTTCCOK 44
DB 54 EYKASNDLYNFRVVDLKSQEQITDSLEYLVEMQWTTCK 96

RESULT 5
Q80ZNS PRELIMINARY; PRT; 141 AA.
ID Q80ZNS;
AC Q80ZNS;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RIKEN cDNA 1700006C19 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strauberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048681; AAH48681.1; -
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C/M.
DR Pfam; PF00031; Cystatin; 1.
DR ProDom; PD001231; Cystatin_C/M; 1.
DR SMART; SM00043; Cy; 1.
DR SMART; SM00043; Cy; 1.
SQ SEQUENCE 141 AA; 16825 MW; C20FA0D8A84951F CRC64;

Query Match 44.9%; Score 122; DB 11; Length 141;
Best Local Similarity 51.2%; Pred. No. 3.6e-08;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYKESDDKXHFRIFRVLKQVQVTDHLEYNLVEMQWTTCCOK 44

```

Db 54 EYKASNDYHFRVVDLKSQEQITDLSLEYLENIARTCK 96

RESULT 6

Q98PX9 PRELIMINARY; PRT; 140 AA.

AC Q98PX9; 01-JUN-2001 (TEMBLrel. 16, Created)
 DT 01-JUN-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Cystatin C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=21010502; PubMed=11144350;
 RA Taupin P.J., Ray J., Fischer W.H., Suh S.T., Hakanson K., Grubb A.,
 RA Gage F.H.;
 RT "FGF-2-Responsive neural stem cell proliferation requires CCG, a novel
 RT autocrine/paracrine cofactor.";
 RL Neuron 28:385-397(2000).
 DR EMBL; AF311741; AAC40283.1; -.
 DR HSSP; P01034; I996.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; Cy; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR CHAIN 21 140 CYSTATIN C.
 FT VARIANT 16 16 A -> G.
 FT VARIANT 84 84 L -> F.
 SQ SEQUENCE 140 AA; 15517 MW; 3A563406D5D8785 CRC64;

Query Match 39.7%; Score 108; DB 11; Length 140;
 Best Local Similarity 43.5%; Pred. No. 2.4e-06;
 Matches 20; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQROVTDHLEHYLNVEMQWTCQK 47
 Db 53 EYKASNDYHSPRIQVVRARQKLVAGVNFPLDVMGRITTCYSQT 98

RESULT 7

Q98SR4 PRELIMINARY; PRT; 112 AA.

AC Q98SR4; 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Cystatin (Fragment).
 OS Acipenser sinensis (Chinese sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 NCBI_TaxID=61970;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two
 RT species of sturgeons.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF334610; AKL6731.1; -.
 DR HSSP; P01038; I990.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR SMART; SM00043; Cy; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.

FT NON TER 1 1
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFE8A08C00 CRC64;

Query Match 35.7%; Score 97; DB 13; Length 112;
 Best Local Similarity 41.9%; Pred. No. 5.4e-05;

Matches 18; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQROVTDHLEHYLNVEMQWTCQK 44
 Db 26 EFNKASNDMYTHRVSKVKQKQVAGIKYIVVQMGRTSCRK 68

RESULT 8

Q98SR3 PRELIMINARY; PRT; 112 AA.

AC Q98SR3; 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Cystatin (Fragment).
 OS Acipenser schrenckii (Ammu sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 NCBI_TaxID=111304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Bai J., Lao H., Ye X., Li Y., Lou J.;
 RT "Molecular cloning and sequence analysis of cystatin cDNA from two
 RT species of sturgeons.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF334611; AKL6732.1; -.
 DR HSSP; P01038; I990.
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; cystatin; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR NON_TER 1 1
 SQ SEQUENCE 112 AA; 12231 MW; 48CEBFE8A08C00 CRC64;

Query Match 35.7%; Score 97; DB 13; Length 112;
 Best Local Similarity 41.9%; Pred. No. 5.4e-05;
 Matches 18; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 2 QYNKESDDKXHFRIFRVLKQROVTDHLEHYLNVEMQWTCQK 44
 Db 26 EFNKASNDMYTHRVSKVKQKQVAGIKYIVVQMGRTSCRK 68

RESULT 9

Q80Y72 PRELIMINARY; PRT; 140 AA.

AC Q80Y72; 01-JUN-2003 (TEMBLrel. 24, Created)
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Cystatin-like 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testicle;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Heien P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toehiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywnicki M.I., Skalska U., Smallie D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testicle;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC048646; AAF48646.1; -;
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR InterPro; IPR003243; Cystatin_C/M.
 DR Pfam; PF00031; Cystatin; 1.
 DR ProDom; PD001231; Cystatin_C/M; 1.
 DR SMART; SM00043; CY; 1.
 DR SMART; SM00043; CY; 1.
 SQ SEQUENCE 140 AA; 16199 MW; 32633E9C4697DA0 CRC64;

Query March 30.9%; Score 84; DB 11; Length 140;
 Best Local Similarity 35.6%; Pred. No. 0.0034;
 Matches 16; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

QY 3 YNKSDDKYHFRIFRLVKVQROVTDHLEHYLNVMQWTTCCOKP 47
 Db 55 YNNAASNDTYLYQVQKLGQGWQLTTGVLYVTYKIGRKCKKNET 99

RESULT 10
 Q12700 PRELIMINARY; PRT; 425 AA.
 ID 012700
 AC 012700;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Exo-1,3-beta-glucanase/1,3-beta-D-glucan glucanohydrolase
 DE (EC 3.2.1.58).
 OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
 OX NCBI_TaxID=27300;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 26077;
 RA Eberlein P., del Rey F., Vazquez De Aldana C.R.;
 RT "Cloning and characterization of 1,3-beta-glucanase-encoding genes
 RT from non-conventional Yeasts.";
 RL Yeast 15:91-109(1999).
 DR EMBL; Z46871; CAAB6951.1; -;
 DR HSSP; P29717; 1CZ1.
 DR GO; GO:0004338; F:glucan 1,3-beta-glucosidase activity; IEA.
 DR GO; GO:0004553; F:hydrolase activity; hydrolyzing O-Glycosyl . . .; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR001547; Glyco_hydro. 5.
 DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 425 AA; 49127 MW; 57F063ABE2FBE274 CRC64;

Query March 30.5%; Score 83; DB 3; Length 425;
 Best Local Similarity 30.0%; Pred. No. 0.015;
 Matches 16; Conservative 10; Mismatches 16; Indels 16; Gaps 2;
 QY 1 DQYNKSDDKYHFRIFRLVKVQROVTDHLEHYLNVMQWTTCCOK 44

Db 269 DQYNNVVVDHHHVVPSAGELQRSIDHTVACNKGMDAKKEYHNNVAGESALTDCAR 328

RESULT 11
 ID 09TY65 PRELIMINARY; PRT; 109 AA.
 AC 09TY65;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Antigen maltose binding protein (Fragment).
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94336252; PubMed=8058358;
 RA Trenholme K.R., Tree T.I., Gillespie A.J., Guderian R., Maizels R.M.,
 RA Bradley J.E.;
 RT "Heterogeneity of IGG antibody responses to cloned Onchocerca volvulus
 RT antigens in microfilaremia positive individuals from Bameraldas
 RT Province, Ecuador.";
 RL Parasite Immunol. 16:201-209(1994).
 DR EMBL; S71364; AAC60509.1; -;
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR NON_TER 1
 FT 1
 SQ SEQUENCE 109 AA; 12701 MW; 3C6B5EF14B082B8 CRC64;

Query March 29.8%; Score 81; DB 5; Length 109;
 Best Local Similarity 34.1%; Pred. No. 0.0066;
 Matches 14; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 4 NKESDDKYHFRIFRLVKVQROVTDHLEHYLNVMQWTTCCOK 44
 Db 24 NEQSNDEYHLMPIKLVSSQVAVGYKMDVQVARSQCK 64

RESULT 12
 ID 09NH95 PRELIMINARY; PRT; 148 AA.
 AC 09NH95;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Ls-cystatin.
 OS Litomosoides sigmodontis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Litomosoides.
 OX NCBI_TaxID=42156;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;
 RT "Characterization and immunological properties of a cysteine protease
 RT inhibitor of the filarial parasite Litomosoides sigmodontis.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF229173; AAF35896.1; -;
 DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
 DR InterPro; IPR000010; Cystatin.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 1.
 DR PROSITE; PS00287; CYSTATIN; 1.
 DR CHAIN 25
 FT 148
 SQ SEQUENCE 148 AA; 16686 MW; 2950AA89CA5339C9 CRC64;

Query March 29.8%; Score 81; DB 5; Length 148;
 Best Local Similarity 34.1%; Pred. No. 0.009;
 Matches 14; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 4 NKESDDKHFRIFRVLKVGROVTDHLEHYLNVEMQWTTCK 44
 Db 62 NQOSNDAYHMPKIKLVKSSQVAGVYKRYEIVARSDCK 102

RESULT 13

ID 016159 PRELIMINARY; PRT; 161 AA.

AC 016159;
 DT 01-JUN-1998 (TEMBLrel. 05, Created)
 DT 01-JUN-1998 (TEMBLrel. 05, Last sequence update)
 DE 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 GN Cystatin-type cysteine proteinase inhibitor.
 OS Brugia malayi (Filarial nematode worm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Brugia.
 OX NCBI_TaxID=6279;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gregory W.F., Blaxter M.L., Maizels R.M.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Gregory W.F., Maizels R.M.;
 RT "Two distinct cystatin-type cysteine protease inhibitors from the
 RT parasitic nematode Brugia malayi";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF015263; AAB69857.1; -;
 DR EMBL, AF171793; AAD51086.1; -;
 DR GO: GO:0004869; F: cysteine protease inhibitor activity; IEA.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00287; CYSTATIN; 1.
 SQ SEQUENCE 161 AA; 18406 MW; 8081351BEE226EB5 CRC64;

Qy 4 NKESDDKHFRIFRVLKVGROVTDHLEHYLNVEMQWTTCK 44

Query Match 29.4%; Score 80; DB 5; Length 161;
 Best Local Similarity 34.1%; Pred. No. 0.013;

Matches 14; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

Db 73 NQOSNDAYHMPKIKLVKSSQVAGVYKRYEIVARSDCK 113

RESULT 14

ID 08K397 PRELIMINARY; PRT; 146 AA.

AC 08K397;
 DT 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DE 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE RIKEN cDNA 1110017E11 gene (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Straussberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC027680; AAH27680.1; -;
 DR GO: GO:0004869; F: cysteine protease inhibitor activity; IEA.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 DR SMART: SM00043; CY; 1.
 FT NON_TER
 SQ SEQUENCE 146 AA; 16380 MW; 9D77BB9A6063A5C4 CRC64;

Query Match 29.0%; Score 79; DB 11; Length 146;
 Best Local Similarity 33.3%; Pred. No. 0.016;

Matches 14; Conservative 13; Mismatches 15; Indels 0; Gaps 0;
 Qy 3 YNKESDDKHFRIFRVLKVGROVTDHLEHYLNVEMQWTTCK 44
 Db 56 YNMGSDSLYYFRDVKYIDAKYQLVAGIKYVLTLDISTECK 97

RESULT 15

ID 09D1B1 PRELIMINARY; PRT; 149 AA.

AC 09D1B1;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DE 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE 1110017E11Rik protein (Cystatin M/E) (Cystatin N homolog).
 GN 1110017E11Rik OR CST6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann M., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pezole G., Quackenbush J.,
 RA Schirml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129S6/SvEvTac; TISSUE=Spleen;
 RA Zeeuwen P.L.J.M., van Vlijmen-Willems I.M.J.J., Hendriks W.,
 RA Merks G.F., Schalkwijk J.;
 RT "A mouse cystatin M/B-null mutation.";
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL, AK003744; BAB22976.1; -;
 DR EMBL, AY093591; AMN1475.1; -;
 DR EMBL, AK078116; BAC37132.1; -;
 DR HSSP; P01036; 1CEW.
 DR MGD; MGI:1920970; Cst6.
 DR GO: GO:0001533; C:confined envelope; IDA.
 DR GO: GO:0008544; P:epidermal differentiation; IMP.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00031; Cystatin; 1.
 SQ SEQUENCE 149 AA; 16796 MW; E713EB920B0FC5 CRC64;

Query Match 29.0%; Score 79; DB 11; Length 149;
 Best Local Similarity 33.3%; Pred. No. 0.017;

Fri Mar 19 07:21:58 2004

us-09-941-314-13.mar18.rspt

Page 6

Matches 14; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Qy 3 YNKSDEKCHFRFRILKQVQRYTDHEIHLNVEMQWTTCK 44
Dd 59 YNMGSLSLYFRDPTQIDAKYQLVAGIKKYITLTIESTECK 100

Search completed: March 18, 2004, 14:21:31
Job time : 29.3611 secs

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

